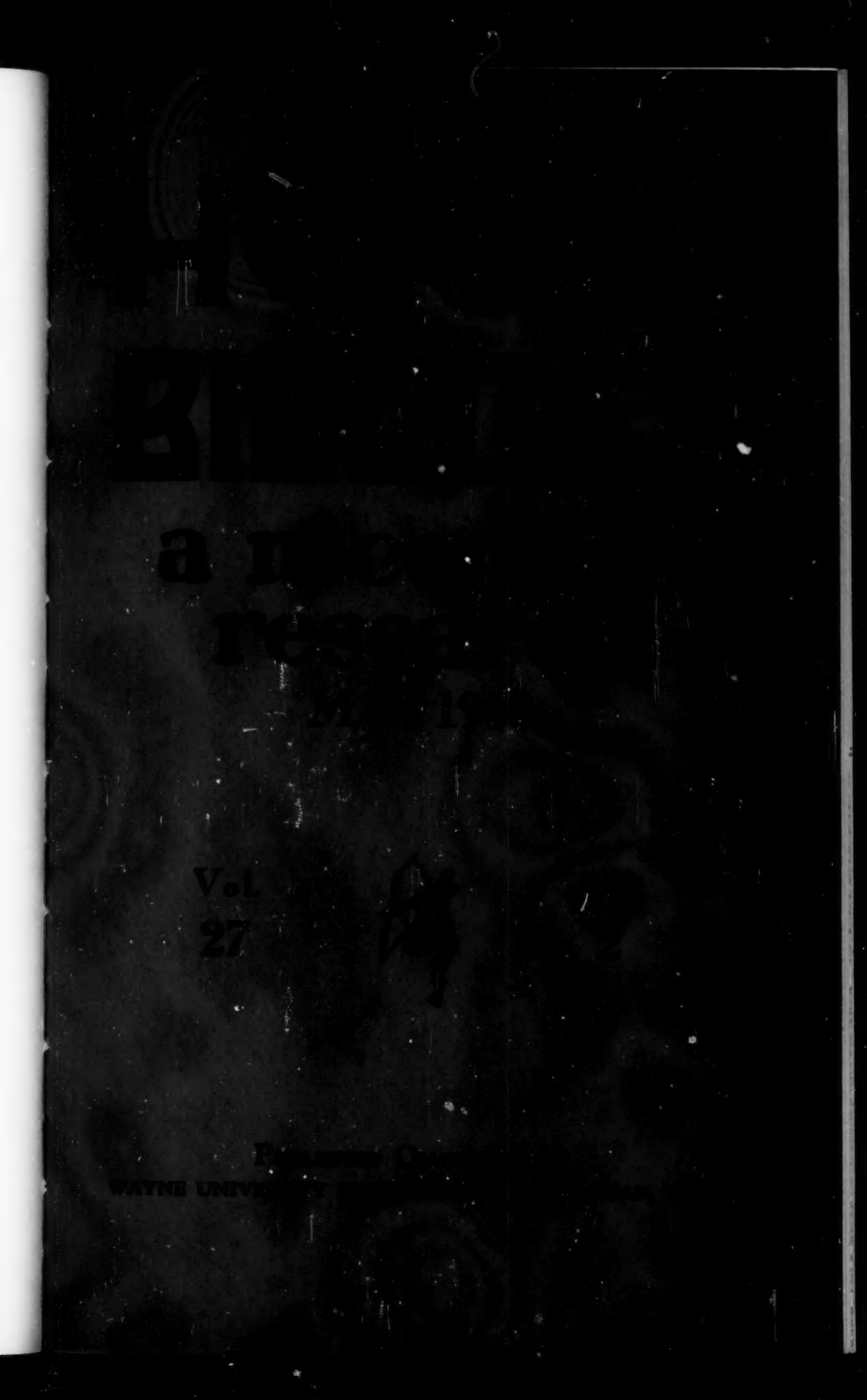

HUMAN BIOLOGY

a record of research

FOUNDED BY RAYMOND PEARL

Editor

GABRIEL W. LASKER

in association with

BENTLEY GLASS — WILLIAM W. HOWELLS

JAMES N. SPUHLER — WILLIAM L. STRAUS, JR.

Review Editor

JOSEF BROŽEK

HUMAN BIOLOGY is published quarterly by Wayne University Press, Detroit 1, Michigan, U. S. A. The advance subscription price is $5.00 a year (foreign postage 35 cents; Canadian 15 cents). Single current issues are $2.00, plus postage.

Subscribers should notify the publishers of change in address at least two weeks before publication of the issue with which the change is to take effect. No claim for non-receipt of an issue will be entertained unless made immediately upon receipt of the next succeeding issue.

Manuscripts and communications regarding editorial matters should be addressed to Editor, Human Biology, Department of Anatomy, Wayne University College of Medicine, 1401 Rivard St., Detroit 7, Michigan.

WAYNE UNIVERSITY PRESS, *Publishers*

4841 Cass Avenue

DETROIT 1, MICHIGAN, U. S. A.

Entered as second class matter at the post office at Detroit, Michigan. Acceptance at the special rate of postage under section 34.40(e) P. L. & R. authorized.

Copyright 1955 by Wayne University Press

HUMAN BIOLOGY
a record of research

MAY, 1955

VOL. 27 No. 2

SKIN-FOLD CHANGES WITH INCREASING OBESITY IN YOUNG AMERICAN MALES *

BY RUSSELL W. NEWMAN

Quartermaster Research and Development Center
Natick, Massachusetts

INTRODUCTION

DURING a recent anthropometric examination of approximately 2,000 young American men, soon after induction into the United States Army and before basic training, skin-fold measurements were obtained. The availability of skin-fold data on such a large series and yet one which was remarkably homogeneous in terms of age and physical selectivity makes it possible to conduct statistical comparisons of subsamples within this group with reasonable validity. The basic description of the skin-fold and related measurements and their demographic distribution will be treated in another report. Two problems are examined here: (1) the question of utilizing skin-fold measurements

* Presented at the annual meeting of the American Association of Physical Anthropologists, Yellow Springs, Ohio, March 28, 1954.

obtained on American Negroes to calculate their body fat; and (2) whether data on large numbers of men of varying degrees of obesity shed any further light on the relationship between subcutaneous fat and total obesity.

It is not the purpose of this paper to review the history and theory of skin-fold measurements as indicators of obesity; this has been covered in a series of recent review articles (Brožek and Keys, '51; Brožek, '53; Keys and Brožek, '53). However, a brief statement of some of the knowns and unknowns is pertinent to the problems presented here. The most interesting aspect of subcutaneous fat as measured by skin-fold thickness is in its relationship with a calculation of total body fat obtained by densiometry. Essentially, this technique consists of a close statistical association between the skin-fold thickness at selected sites and body fat (as a percent of weight) calculated from the average density (weight divided by body volume). Brožek and Keys ('51) have demonstrated such a close relationship for 133 University of Minnesota college men that a multiple regression equation based on the skin-fold thickness at the chest, arm, and abdomen can be used to calculate body fat with a high degree of reliability. There are a number of limitations inherent in the use of this equation, only two of which will be discussed here: (1) since this equation was based on White college students, can it be applied equally to other racial groups, in this case American Negroes? (2) since it is a single multiple regression equation which fits all young men whether lean or obese, it assumes that the skin-fold thicknesses at all sites represented in the equation increase linearly with increasing obesity. Is this true when a large sample is analyzed? Unfortunately, one vital portion of the data on the Army men, a reliable and independent measure of body volume from which average density could be calculated, was not available. None of the densiometric techniques now available or expected in the near future (*e. g.*, underwater weighing, air displacement, gas dilution), could handle such large numbers of men in the time available. This lack of an independent method of calculating body fat limits the analyses to the observed data, skin-fold thickness, body fat calculated by the Brožek and Keys' equation, and total body weight.

Some of the pertinent elements of the measuring technique are reported here because the increasing number of measuring instruments and skin-fold sites in use might otherwise eventually result in the grouping of basically inapplicable data. Five skin-fold sites were

measured on these subjects: chest—above and to the right of the right nipple; arm—on the back side, half-way up the upper arm; back—immediately below the right scapula; abdomen—to the left of the navel; and knee—above the right kneecap. The subjects were measured in an upright position by two technicians and their average value was used. Each investigator measured every skin-fold site at least three times before proceeding to the next area. The calipers used were a slightly modified version of the Minnesota type skin-caliper similar to that described by Brožek ('52) and illustrated by Newman ('52). The caliper points had a diameter of 6 mm with a surface pressure of approximately 5 g per sq mm.

The series used here were 1,702 White males and 292 American Negro males. This latter group will be referred to subsequently as Negroid to acknowledge the obvious racial admixture which was present in this sample. The average age of these men was 20.7 years with a relatively restricted age range, 17 to 28 years. The effect of physical selection on these men is unknown, but it probably deviates no more from a truly random sample of young American men than a series based on volunteer college students.

NEGROID-WHITE COMPARISONS

Any detailed comparision of White and Negroid skin-folds is complicated by the fact that the Negroids are superficially far leaner than the Whites. For example, the average Negroid body fat calculated from Brožek and Keys' regression equation was 3.9%, approximately half the comparable white value of 6.8%. This preponderance of small skin-fold values in the Negroids results in a lack of individuals whose skin-fold values can be compared with the obese White subjects. This paper is not concerned with why the Negroids show lower skin-fold values but whether their subcutaneous fat is relatively rather than absolutely comparable to Whites in terms of gross measured quantities.

One method of investigating the relationships between skin-fold sites, irrespective of the amount present, is to calculate correlation coefficients for the 5 areas. The 10 possible coefficients are shown in table 1 for both racial groups. It is apparent that Whites and Negroids show a similar arrangement of coefficients in bodily areas; the chest, back, and abdomen skin-folds are highly intercorrelated with each other and less correlated with the extremities (arm and knee) which form another correlated set. The two matrices of coefficients are very close in magni-

tude, but the Negroid coefficients are slightly lower in value. There does not seem to be any readily apparent reason for this difference unless it is the effect of racial mixture appearing as heterogeneity in the coefficients.

A statistical test of significant difference between the comparable White and Negroid correlation coefficients was applied to the 10 pairs given in table 1. Since the numbers of men involved for each site relationship were constant, a common significance level of t of 1.96 at the 5% level and 2.58 at the 1% level were used. The results are shown below in table 2. It is difficult to evaluate such a table where two

TABLE 1

Correlation Coefficients between Skin-fold Sites of Measurement

	ARM	BACK	KNEE	ABDOMEN
WHITES				
Chest	.79	.90	.73	.90
Arm		.78	.83	.81
Back			.74	.87
Knee				.75
NEGROIDS				
Chest	.79	.80	.67	.86
Arm		.74	.82	.80
Back			.68	.86
Knee				.72

relationships show significant difference and 8 do not, but it would appear that the skin-fold relationships between the bodily areas are at least roughly similar in these groups. A possible explanation for the difference in chest-abdomen coefficients may be brought out in the next section.

Lacking any measure of body volume, it was not possible to construct regression equations between skin-fold values and calculated body fat for these two groups and to compare them. Furthermore, it was not possible to throw much light on the question of racial differences in body fat from a regression equation based on Whites alone. Because of this postulate, the samples were divided into increments of apparent obesity on the best evidence available, a summation of the measured skin-folds. The term "apparent obesity" is used for two reasons: (1) It must be assumed that the two principal fat depots, superficial (subcutaneous) and deep (largely visceral), increase in roughly equal proportion with increasing obesity. This has been shown in Whites by the close corre-

spondence between skin-folds and tissue density, but it remains an assumption for the Negroids. (2) It is apparent from Brožek and Keys' ('51) prediction equation between skin-fold thickness and body-fat from densiometry that the various sites do not have equal weight in this association, for example, abdomen, chest, and arm have relative values of 15, 39, and 46%, respectively. No differential weightings could be used with this Army data without prejudicing the results, and it is believed that no large error has been introduced by the simple addition of the skin-fold measurements into a summation value which must roughly align the men into increments of increasing obesity.

TABLE 2

Comparison of White and Negroid Values of r

SITE	"t" VALUE	SIGNIFICANCE
Chest-arm	0.16	None
Chest-back	5.47	1%
Chest-knee	1.88	None
Chest-abdomen	2.81	1%
Arm-back	1.41	None
Arm-knee	0.62	None
Arm-abdomen	0.47	None
Back-knee	1.72	None
Back-abdomen	1.25	None
Knee-abdomen	0.94	None

The White and Negroid samples were divided into 5 mm intervals based on the sum of abdominal, arm, and chest skin-folds. The mean value for each of the three skin-fold sites for all men in each summation interval was calculated and plotted in figure 1. The entire sample of 1,702 Whites and 292 Negroids is not represented in this figure which shows only those individuals whose combined skin-fold value totaled 45 mm or less. Above this value there were not enough Negroids to avoid erratic sampling errors. The actual numbers of men represented by the plotted values in figure 1 are given in appendix A.

The lines of mean skin-fold values in this figure closely approximate regression lines and may be tested statistically for significant differences in slope. This was accomplished for each racial pair and the lines were found to be not significantly different. It is apparent, however, that the Negroid has low skin-fold values in chest and arm and accompanying high abdominal skin-fold thickness as contrasted with the Whites. These differences were visible at the time of measurement. The Negroids were

characterized by well-developed and well-defined musculature in the torso and upper extremities, and musculature could not be well-defined when covered by a thick layer of subcutaneous fat. Hooton ('48) mentioned this when comparing body build photographs in 3,000 Negroids and 30,000 Whites. He said that "the shoulders of the Negro customarily show relatively greater muscular development than is found in Whites.

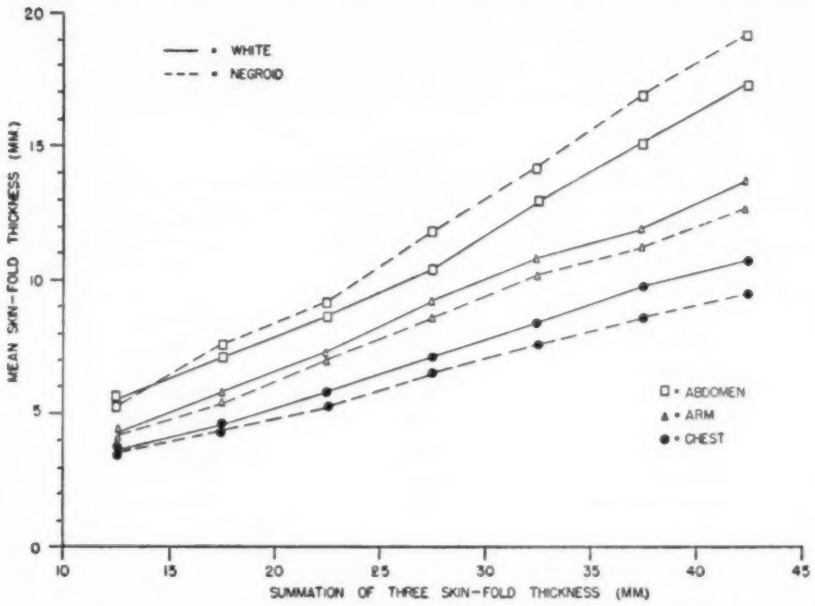

FIG. 1. COMPARISON OF THREE WHITE AND THREE NEGROID SKIN-FOLDS.

The Trapezii and Deltoids are characteristically well developed, as are also the pectoral muscles, the abdominal muscles, and those of the back. Thigh muscles are also generally good."

The minimal values for men who have 10 to 15 mm total for the three skin-fold sites are remarkably consistent between the racial groups. Two points may be inferred from this: (1) the average value must be very close to fat-free, doubled skin-fold thickness; (2) it suggests that there is no measurable difference in fat-free skin-fold thickness between Negroids and Whites, at least at these particular sites. Tanner ('51) records three millimeters as the minimum skin-fold thickness found in approximately 400 British students. No skin-fold reading of a lower magnitude was recorded for the 2,000 men reported here.

It is possible to estimate the amount of body fat by substituting the mean values given in figure 1 for any given interval in the Brožek and Keys' ('51) prediction equation. This assumes that the overall subcutaneous fat to average body density relationship is basically similar for the two groups, and further that the Negroid superficial versus deep fat ratio is the same as that found in the White subjects. A calculation of percent body fat using the mean skin-fold values, rounded to whole millimeters, for the largest interval used in figure 1 (40 to 45 mm) gave a White value of 8.4% and a Negroid value of 7.9%. One half of one per cent is a very small difference for a measurement as full of potential pitfalls as the estimation of body fat.

The first question asked in this study, whether skin-fold measurements on American Negroes can be converted to estimates of per cent body-fat from relationships established on White subjects, cannot be answered definitively from the data presented here, but it appears that no great error would result from such a practice. The relationships shown in table 1 and figure 1 are considered sufficiently similar to warrant using the total series of 1,994 men in examination of the next question, whether data on a large group of young men can show any new relationships between skin-folds and obesity.

MEAN VALUES FOR THE COMBINED SERIES

The use of the combined series, which now might be categorized as young American male, offers a much larger range of obesity than could be shown in figure 1 because of the inclusion of the fatter White subjects. The average values at all 5 measured skin-points have been plotted for all the summation of total skin-fold thickness intervals in which 10 or more individuals were represented in figure 2. The largest summation interval used in this figure (130 to 139 mm) represents approximately 20% body fat. The actual number of individuals used in the calculation of the mean skin-fold values for each summation interval in figure 2 are given in appendix A.

There are several relationships to be seen in this figure. The most striking is the curvilinear relationship of the chest skin-fold to increasing total thickness. If this trend continues into the range of even greater obesity, it becomes second only to the abdominal site in importance. Both linear and second degree curve regression equations were calculated on the chest mean values and the standard error of the curvilinear equation was significantly smaller than that of the linear equation. This

curvilinearity may be caused by some peculiarity of this series, and, in any case, is not of much practical importance.

Even more interesting is the difference in the slope of the lines across the range of obesity represented here. If a line perpendicular to the abscissa were drawn about midway on the summation scale, in figure 2, it could be considered to divide the series into those leaner than average

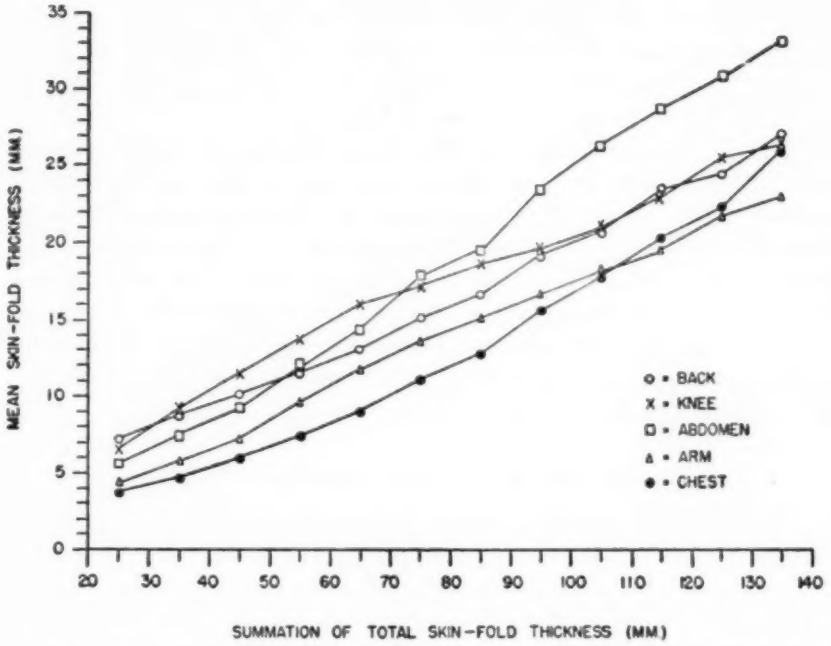

FIG. 2. CHANGES IN MEAN SKIN-FOLD VALUES WITH INCREASING TOTAL THICKNESS FOR TOTAL SERIES (2000 MALES).

on the left and those fatter than average on the right. The fatter group on the right half of the figure shows considerably different slopes than characterize the leaner group. The marked inclination of the abdominal line is readily apparent; this trend is present on the left half but is not as distinctive in terms of millimeters of skin-fold thickness. It is also apparent that the chest slope is roughly parallel to that of abdomen and both are at odds with the slope of arm and knee, while the slope of back skin-fold appears to have an intermediate position. This probably should be interpreted as illustrative of a differential development of subcutaneous fat with increasing obesity. The torso, especially on the

anterior aspect, has a larger capacity to store subcutaneous tissue and becomes increasingly important in fatter individuals, while the extremities are limited in the physical space available for this tissue and, therefore, grow less important. If this situation also occurs when a large sample of men are arranged in ascending order of body density,

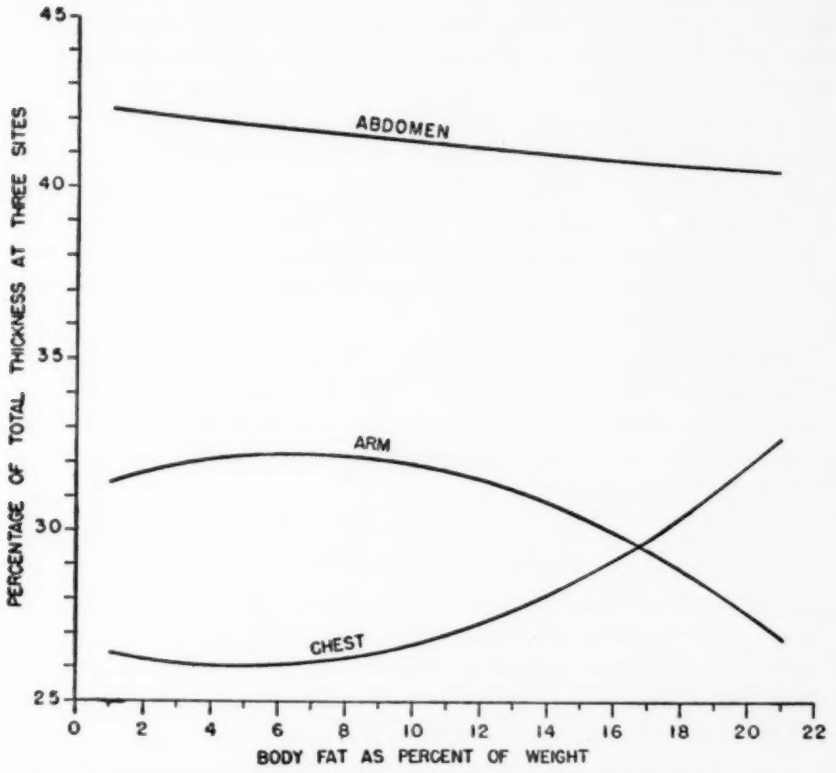

FIG. 3. PERCENTAGE RELATIONSHIPS OF THREE SKIN-FOLD THICKNESSES WITH INCREASING OBESITY (TOTAL SERIES).

the use of any one multiple regression equation for calculating percent body fat cannot be entirely satisfactory. It probably makes little difference in men of average fat values, but it must decrease the accuracy of prediction on the very lean and the very fat.

One additional graphic method of representing the relationships of the various skin-fold sites with increasing obesity has been used in figure 3. In this graph the mean skin-fold values at abdomen, arm, and chest

have been expressed as percentages of the total thickness on the ordinate axis. These percentages have been plotted against calculated percent body fat. Since the body fat data were derived from a multiple regression equation of these three skin-folds, the values of fat and skin-folds cannot be independent of each other.

A plot of percentage values such as shown in figure 3 maximizes any change in proportions at the three sites because as one site value rises some other must decline to preserve a totality of 100% for each interval. Only the calculated lines of best fit have been used in this figure to achieve a simplified graphic effect. The data from which the lines were calculated are given in appendix A. The lines of best fit for arm and chest are obviously curvilinear in this figure. Linear and second degree curve regression equations were calculated from the mean percentage values for each interval and the standard error of the arm and chest second degree equations were significantly smaller than the standard errors of the linear equations.

The arm and chest lines in this figure form a reciprocal pair, similar to the crossover shown in figure 2. This presumably should be interpreted in the same manner, namely, that with increasing obesity subcutaneous fat is deposited in relatively greater quantities in the torso than in the extremities.

SUMMARY AND CONCLUSIONS

Two problems involving skin-fold measurements on a large series of young American men have been examined in this paper. The results cannot be considered definitive because no independent measure of average body density was available; the conclusions arrived at using the sum of the skin-folds can only be termed indicative.

An analysis of Negro and White skin-fold relationships by comparing correlation coefficients between skin-fold sites and average skin-fold values for increments of obesity showed small and not statistically significant differences between these racial groups. It was concluded that the present data do not show any reason why body fat on Negroes cannot be calculated from the published prediction equation based on White subjects. However, the comparison was limited by the scarcity of obese Negroes, and it is possible that the observed Negro deficiency of fat over the pectoral and triceps regions might become practically significant if very obese Negroes were compared to their White counter-

parts. This racial difference in the regional distribution of subcutaneous fat is an intriguing problem which should repay further investigation.

A graphic comparison of the various skin-fold sites in the combined series of 2,000 men for presumed increments of increasing obesity indicated a change in the proportional representation of the average site values between the lean and fat subjects. The anterior torso areas, chest and abdomen, showed a rise in relative importance with increasing obesity at the expense of the extremity areas, *i. e.,* the arm and knee. This was interpreted to be at least partially the result of physical space limitations in the extremities. It may form a source of error in predicting fat on all men of a comparable age group with a single prediction equation.

This paper has only touched on a few of the interesting possibilities inherent in the skin-fold measurements of these 2,000 young men. Its major purpose was to validate, if possible, the calculation of body fat on the 361 Negroids for use in other analyses. It is believed that this has been accomplished, and a new and intriguing field of anthropometric racial investigation can be developed through the skin-fold measurement technique.

LITERATURE CITED

BROŽEK, J. 1952 Changes of body composition in man during maturity and their nutritional implications. Fed. Proc. *11*: 784-793.

——— 1953 Measuring nutriture. Am. J. Phys. Anthrop. n. s., *11*: 147-180.

BROŽEK, J., AND A. KEYS 1951 The evaluation of leanness-fatness in man, norms and interrelationships. Brit. J. Nutr. *5*: 194-206.

HOOTON, E. A. 1948 Body build in relation to military function in a sample of the United States Army. Prepared for the Office of The Quartermaster General. From the Department of Anthropology, Harvard University, Cambridge, Mass.

KEYS, A., AND J. BROŽEK 1953 Body fat in adult man. Physiol. Rev. *33*: 245-325.

NEWMAN, R. W. 1952 Measurement of body fat in stress situations. Office of The Quartermaster General, Research and Development Division, Environmental Protection Branch Report No. 193.

TANNER, J. M. 1951 The relation between serum cholesterol and physique in healthy young men. J. Physiol. *115*: 371-390.

APPENDIX A

Additional Data for Figures 1, 2, and 3

FIGURE 1			FIGURE 2			FIGURE 3				
mm Skin-Fold	Frequency White	Negroid	mm Skin-Fold	Frequency White	Negroid	Per cent Body Fat	Number	Mean Percentage of Total Thickness Chest	Arm	Abdomen
10-14	23	12	25-29	16	8	0-1	74	26.7	30.3	43.0
15-19	201	83	30-34	75	24	2-3	554	25.8	32.0	42.2
20-24	345	88	35-39	175	56	4-5	490	26.1	33.1	40.8
25-29	266	45	40-44	203	56	6-7	321	26.1	32.9	41.0
30-34	230	21	45-49	181	39	8-9	194	25.6	33.0	41.4
35-39	157	18	50-54	172	25	10-11	118	27.4	31.4	41.2
40-44	111	10	55-59	151	19	12-13	76	27.7	30.7	41.6
			60-64	150	19	14-15	58	29.5	28.8	41.7
			65-69	92	13	16-17	36	29.1	29.4	41.5
			70-74	75	10	18-19	26	31.4	28.2	40.4
						20-21	22	32.5	27.6	39.9

AGE DIFFERENCES IN THORACIC FORM AS INDICATED BY THORACIC INDEX

BY EIJI TAKAHASHI AND HIROSHI ATSUMI

Department of Hygiene and Public Health, Hirosaki University, Japan

MARTIN ('28) and Mollison ('33) explain the flatter build of the human thorax, compared with that of other primates and mammals, by the change to upright position. This phylogenetic tendency to a flatter thorax is also seen in ontogenesis. The human thorax becomes flatter during the fetal and postnatal growth periods (Martin, '28). We have followed the transformation of the form of the thorax in Japanese in all the age periods of postnatal life.

The present study is based on body measurements of 9369 inhabitants of the Tōhoku area, especially Tsugara district of Aomori Prefecture in Japan. These individuals, from birth to 85 years of age, were measured for stature, weight, chest depth, breadth and girth, and some other measurements. Chest dimensions were taken at the end of expiration of tidal air. The percentage of chest depth to chest breadth is taken as the thoracic index, although Martin used the inverse.

The subjects were mostly rural inhabitants. Children from 6 to 11 years old were primary school pupils and those from 12 to 14 attended middle school; these are periods of compulsory education in Japan. The subjects from 15 to 17 were high school pupils and almost all those from 18 to 20 were also students. The adult group consisted of farmers, teachers, officials, and service workers, and the people over 60 years of age had had various careers and were mostly inmates of an institution for the aged.

FINDINGS

The findings are shown in table 1 and in figure 1. The average thoracic index of the newborn is 84 in both sexes. After birth the index decreases rapidly until in 4-year old children it is about 72 in boys and

TABLE 1

Stature and Thoracic Index by Sex and Age
(means and standard errors)

	MALE			FEMALE		
Age	Number	Standing Height (cm)	Thoracic Index	Number	Standing Height (cm)	Thoracic Index
0–6 mos	74	63.4±0.49	83.88±1.37	71	62.1±0.47	83.72±0.94
6–12	88	69.0±0.35	80.28±0.76	86	68.2±0.44	79.80±0.68
1–2 yrs	142	76.3±0.38	78.24±0.44	131	74.3±0.33	77.96±0.60
2–3	82	85.2±0.51	76.72±0.57	79	82.1±0.53	75.40±0.50
3–4	83	91.2±0.51	74.02±0.45	77	90.7±0.43	72.61±0.66
4	33	100.2±0.92	72.24±0.79	31	98.5±0.98	71.00±0.84
5	36	104.5±0.73	71.88±0.66	36	103.5±0.76	70.70±0.82
6	115	109.1±0.36	72.14±0.37	132	108.8±0.36	70.68±0.39
7	175	114.1±0.32	72.16±0.33	166	113.6±0.37	70.38±0.41
8	181	119.9±0.37	71.46±0.16	185	118.3±0.35	70.05±0.32
9	181	125.1±0.37	71.52±0.33	174	123.9±0.37	70.34±0.36
10	212	129.4±0.34	71.12±0.30	180	128.6±0.43	69.76±0.35
11	185	133.4±0.36	71.76±0.34	189	133.8±0.47	69.18±0.32
12	168	138.5±0.49	71.90±0.33	159	141.6±0.55	68.98±0.37
13	321	144.6±0.42	71.64±0.25	272	145.1±0.39	71.06±0.28
14	258	151.4±0.46	72.16±0.26	286	148.8±0.32	71.48±0.35
15	506	158.2±0.31	70.13±0.24	389	151.2±0.25	71.21±0.27
16	434	162.0±0.26	69.20±0.25	275	152.9±0.27	71.28±0.34
17	377	163.2±0.28	68.98±0.23	198	153.5±0.33	71.62±0.40
18	322	164.4±0.28	70.08±0.30	184	154.3±0.39	72.24±0.36
19	289	165.5±0.31	68.92±0.30	89	154.0±0.51	72.02±0.56
20	191	165.1±0.38	68.80±0.43	43	153.2±0.64	71.16±0.94
15–19*	91	161.3±0.68	69.20±0.50	66	151.3±0.61	70.64±0.48
20–24	170	164.9±0.39	68.52±0.44	123	152.1±0.48	72.00±0.60
25–29	127	164.7±0.53	67.92±0.59	87	151.6±0.64	72.76±0.60
30–39	179	162.4±0.48	69.48±0.39	102	149.2±0.57	72.98±0.50
40–49	158	162.3±0.54	70.24±0.47	91	148.8±0.64	74.62±0.61
50–59	141	159.6±0.71	72.18±0.53	67	145.9±0.91	77.02±0.71
60–69	18	(154.0±1.60)	76.67±2.06	17	(144.4±1.57)	82.05±2.43
70 & over	26	(152.5±1.67)	80.38±1.68	30	(140.2±1.61)	83.33±1.78

N. B. Subjects from 6 to 11 years are primary school children, from 12 to 14 pupils of middle school, from 15 to 17 pupils of high school and from 18 to 20 also mostly students.

* 15–19 year old youths who do not attend school.

71 in girls. It seems that the decrease is greater in girls. The average thoracic index continues to decline in girls 4 to 12 years of age when it is about 69, significantly lower than in boys of the same age. In other studies of Japanese (figure 2) and in American Negroes (figure 3) the thorax is also flatter in females of these ages.

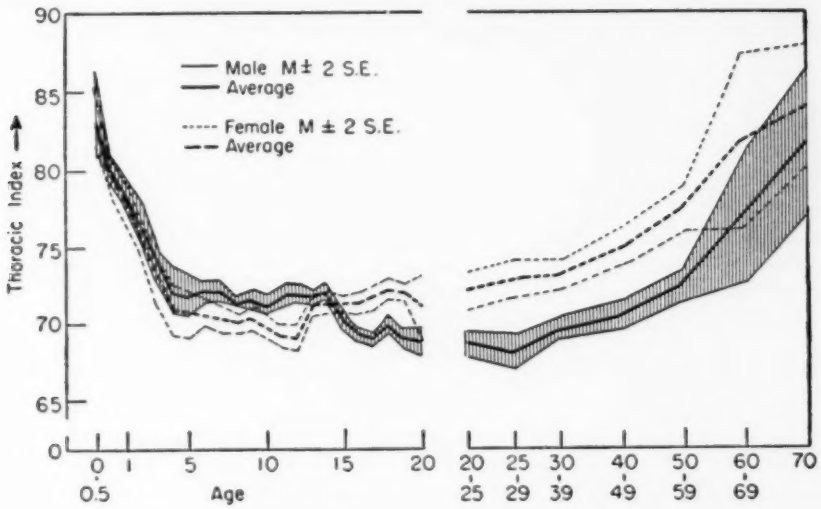

FIG. 1. THORACIC INDEX CURVE OF JAPANESE IN TSUGARA DISTRICT.

The thoracic index of girls age 13 appears to be suddenly larger (to about 71) while the index of the males goes down from 15 to 25–29 years of age (to about 68) so that the male chest is flatter from between 14 and 15 years of age onward. From about the age of 25 to 29 chest index begins to be larger again, increasingly so after the 40–49 year old period and this is apparent in both sexes. In the age class of 70 and over (really 70–85 years of age) the index is again over 80 in both sexes as in the newborn.

DISCUSSION

The chest depth and breadth of Japanese have also been measured by some other authors (figure 2), but not for all age classes. Differences in the thoracic index reported by different authors seems to be partly explained by differences in the province, occupation or rural-urban nature of the population studied; it may also represent differences in technique involving different phases of respiration or a different level

of the chest in the measurements of different authors. In any case, all authors report the same tendencies as to sex and age differences. The other studies of Japanese accord with ours in age groups up to 11 or 12

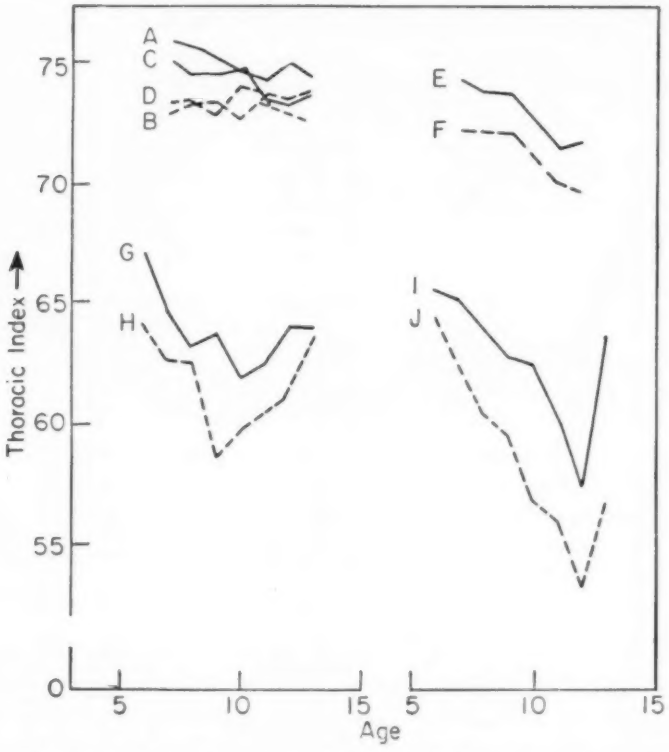

FIG. 2. THORACIC INDEX CURVE OF JAPANESE SCHOOL CHILDREN BY OTHER AUTHORS.

Solid line: male; dotted line: female

(a) (b) (Urban) MINAYOSHI, T. 1928, Nippon Gakkō Eisei (in Japanese), *10*: 721-742.

(c) (d) (Rural) MINAYOSHI, T. *Ibid.*

(e) (f) SHIMA, S. 1931, Nippon Gakkō Eisei (in Japanese), *19*: 199-204.

(g) (h) DŌSONO, S. 1946, Essays of Anthropology, Human Genetics and Constitution (in Japanese), Keiō Gizyuku Univ., *6*: 1-115.

(i) (j) INAGAKI, T. AND A. NAKAZAWA 1947, *Ibid.*, *11*: 55-165.

years of age, but Martin's ('28) data on Negroes is most similar to ours in all age classes. All groups seem to show that the rounder chest of the newborn becomes increasingly flatter until about 4 years of age,

that before puberty girls have relatively flatter chests than boys and that after puberty girls' chests become markedly rounder. After 14 years of age the period of compulsory education ends and few children of the less privileged classes attend school; there is, therefore, a selective factor in our series. Nevertheless, there seems to be a real tendency to further flattening in the chest of males until 25–29 years. The increased flattening of the chest in successively older persons up to this age seems

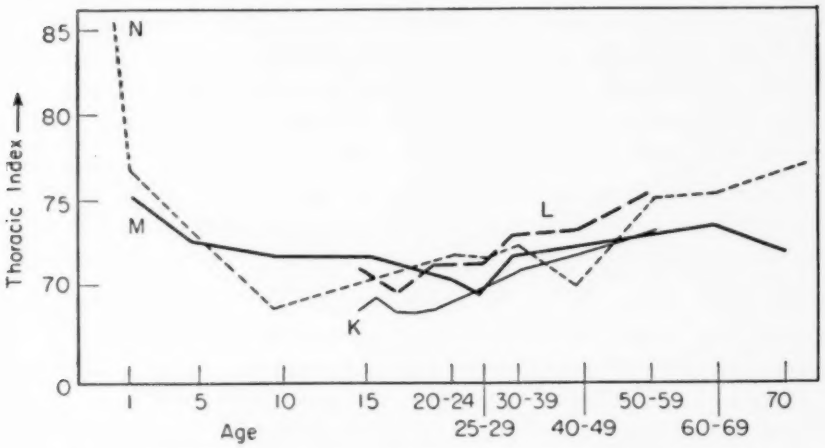

FIG. 3. THORACIC INDEX CURVE BY SEX AND AGE BY OTHER AUTHORS. Solid line: male; dotted line: female.

(k) (l) Japanese prisoners, DOHI, N. 1949, Essays of Anthropology, Human Genetics and Constitution (in Japanese), Keiō Gizyuku Univ., *14*: 15-206.

(m) (n) American Negroes, MARTIN, R. 1928, Lehrbuch der Anthropologie, Gustav Fischer, Jena.

to be a developmental phenomenon. Although chest depth also increases, the increase in chest breadth is more rapid. Martin ('28) says that chest breadth always increases at the cost of chest depth. Chest breadth shows a significant positive correlation with stature, but chest depth does not (table 2). The greater relative depth of the female chest after puberty seems to represent a real sex difference; it is, of course, at this time that there is a rapid development of the female breasts.

Many authors agree with us in noting a cessation of growth in chest breadth after 30–39 years of age but a continued slow increase in chest depth. The thorax thus becomes rounder, but this may be a regressive rather than a developmental phenomenon.

In order to analyze the thoracic index further, we calculated separately that of adults of different occupation groups (table 3). In both males and females the lowest index is found in teachers and progressively higher indices are found in officials, service workers and farmers respectively. This same tendency is seen in the different age periods considered separately, and, conversely, the tendency to an increase in the thoracic index with age occurs in each occupational group.

TABLE 2

Coefficients of Correlation between Diameters of Chest, Thoracic Index, Relative Transverse Area of Chest, and Other Physical Measurements of Adults

	CHEST DEPTH		CHEST BREADTH		THORACIC INDEX		RELATIVE TRANSVERSE AREA OF CHEST	
	Male	Female	Male	Female	Male	Female	Male	Female
Standing height	0.040	0.017	0.245**	0.130*	—0.184**	—0.121*	0.106**	0.165**
Sitting height	0.022	0.025	0.250**	0.158**	—0.193**	—0.122*	0.104**	0.137**
Girth of chest	0.583**	0.622**	0.628**	0.600**	0.087*	0.110*	0.695**	0.679**
Body weight	0.544**	0.563**	0.616**	0.574**	0.009	0.198**	0.628**	0.646**
Pelidisi	0.494**	0.504**	0.390**	0.421**	0.193**	0.190**	0.567**	0.809**
Vital capacity	0.130**	0.167**	0.372**	0.327**	—0.037	—0.026	0.311**	0.276**
Strength of back muscles	0.110*	0.149*	0.230**	0.247**	—0.036	—0.003	0.259**	0.419**

N. B. Number of measured adults: 656 males and 350 females.

* Over the level of significance of P = 0.05.

** Over the level of significance of P = 0.01.

The relative thickening of the chest in older age may be caused by the increased spinal curvature or by the decrease in height of the thorax which follows the loss of fluid and compression of the *disci intervertebrales.*

The thorax is composed of the spine (vertebrae), ribs and sternum. The spine is the most stable of these three elements and the motility of the ribs increases from the vertebral to the sternal end. Chest breadth is less affected by movement than is chest depth. The difference between measurements on inspiration and on expiration is less for chest breadth than for chest depth (table 4). Furthermore, chest depth tends

TABLE 3

Chest Depth, Breadth and Thoracic Index by Sex and Occupational Group (means and standard errors)

SEX	AGE CLASS	TEACHERS	OFFICIALS	SERVICE WORKERS	FARMERS
Male					
	20–29	18.03±0.11	18.15±0.15	18.44±0.23	18.98±0.21
Chest depth	30–49	18.64±0.12	18.63±0.16	19.10±0.30	19.15±0.21
(cm)	50–59	18.99±0.25	19.32±0.24	19.30±0.25	19.40±0.28
	Average	18.55±0.06	18.70±0.07	18.95±0.07	19.16±0.08
	20–29	26.63±0.12	26.45±0.15	26.29±0.29	27.11±0.21
Chest breadth	30–49	26.99±0.13	26.88±0.16	26.77±0.27	27.06±0.20
(cm)	50–59	27.04±0.26	26.94±0.24	26.70±0.26	26.40±0.34
	Average	26.89±0.04	26.76±0.05	26.59±0.02	26.86±0.07
	20–29	68.08±0.49	69.84±0.05	69.88±0.91	70.14±1.04
Thoracic index	30–49	69.20±0.52	69.64±0.66	70.92±0.94	71.02±0.97
	50–59	70.06±1.12	72.00±1.11	72.46±1.15	73.54±1.16
	Average	69.11±0.28	70.49±0.34	71.09±0.37	71.64±0.38
Female					
	20–29	16.91±0.14	16.53±0.20	17.80±0.17	17.94±0.25
Chest depth	30–49	17.00±0.18	17.50±0.18	18.21±0.17	18.52±0.21
(cm)	50–59	——	——	(18.47±0.26)	(18.92±0.25)
	Average	16.96±0.08	17.02±0.10	18.02±0.09*	18.23±0.09*
	20–29	23.72±0.13	23.41±0.21	24.44±0.12	24.56±0.28
Chest breadth	30–49	23.28±0.20	23.19±0.24	24.48±0.15	24.34±0.31
(cm)	50–59	——	——	(24.12±0.27)	(24.30±0.30)
	Average	23.50±0.08	23.30±0.10	24.46±0.07*	24.45±0.08*
	20–29	71.46±0.52	70.42±0.72	72.76±0.73	73.16±1.05
Thoracic index	30–49	72.88±0.76	74.24±0.70	74.50±0.61	76.10±0.92
	50–59	——	——	(76.58±1.11)	(77.91±1.05)
	Average	72.12±0.39	72.33±0.49	73.57±0.38*	74.60±0.40*

* Data for the age class 50–59 are excluded.

to be more variable; for example, the difference in mean chest breadth between male teachers and farmers (as shown in table 3) is slight and not significant but the difference in chest depth is statistically significant.

As sitting height decreases with age and thoracic contents do not decrease correspondingly, there must be an adjustment in the breadth

TABLE 4

Comparison of the Difference between Chest Depth and Breadth in Deep Respiration

AGE CLASSES	NUMBER MEASURED	IN DEEP INSPIRATION (cm) (I)	IN DEEP EXPIRATION (cm) (E)	DIFFERENCE (cm) (I—E)	$\frac{I-E}{\frac{1}{2}(I+E)} \times 100$
Chest Depth					
20–29	315	19.46	17.41	2.05	11.12
30–39	130	19.64	17.50	2.14	11.52
40–49	59	20.14	17.92	2.22	11.66
50–59	31	19.77	17.94	1.83	9.70
Average	535	19.59	17.53	2.06	10.92
Chest Breadth					
20–29	315	26.49	24.91	1.58	6.14
30–39	130	26.67	24.95	1.72	6.66
40–49	59	26.56	25.03	1.53	5.93
50–59	31	25.71	24.51	1.20	4.77
Average	535	26.47	24.91	1.56	6.07

N. B. Calculated from the data of DOHI on male prisoners (Essays of Anthropology, Human Genetics and Constitution, edited by Dr. Taniguchi, Vol. 14, 1949, Keiō Gizyuku University, Japan. In Japanese).

and depth of the thorax. The ribs themselves and the thoracic girth increase little, if at all, and the increase is in depth rather than in breadth. The increased area of the thoracic cross section is achieved by the change from an ellipse towards a circle.

In Japan there are many people with marked spinal curvature which is not caused by any disease, but can be ascribed to the way of life and work—for instance, rice farming by the Japanese method, sewing and sitting on floor mats in the Japanese manner. Table 5 shows the correlation between senile curvature of the spine and roundness of the thorax. The spinal curve in the sagittal plane from the seventh cervical

to the fourth lumbar vertebra was copied with a lead and plastic foc drawing curve ruler. The index of spinal curvature is expressed as the ratio of the length of the perpendicular (height of the curve) to the length of the cord. We calculated the coefficients of correlation between age, stature, thoracic index and index of spinal curvature on the basis of investigations of 58 senile persons housed in an asylum. The correlations of all of these factors are statistically significant except for age with thoracic index. The correlation between thoracic index and spinal curvature is especially marked and is demonstrable in the second

TABLE 5

Correlation Coefficients among Age (1), *Stature* (2), *Index of Spinal Curvature* (3) *and Thoracic Index* (4) *in Old People*

$r_{12} = -.304$	$r_{13} = +.398$	$r_{14} = +.244$
$r_{12.3} = -.152$	$r_{13.2} = +.306$	$r_{14.2} = +.145$
$r_{12.4} = -.235$	$r_{13.4} = +.326$	$r_{14.3} = +.391$
$r_{12.34} = -.114$	$r_{13.24} = +.273$	$r_{14.23} = +.283$
$r_{23} = -.453$	$r_{24} = -.380$	$r_{34} = +.554$
$r_{23.1} = -.380$	$r_{24.1} = -.331$	$r_{34.1} = +.513$
$r_{23.4} = -.314$	$r_{24.3} = -.174$	$r_{34.2} = +.463$
$r_{23.14} = -.258$	$r_{24.13} = -.136$	$r_{34.12} = +.498$

order; that is, it is siginficant even when the influences of age and stature are excluded. In senile persons with markedly crooked spines the chest depth is sometimes greater than the chest breadth as in most quadrupedal mammalia. These extremely bent individuals walk with the aid of a stick holding the upper part of the body parallel to the ground in much the same manner as a quadruped. If the influence of the gravity of the thoracic contents is deduced to explain the form of the chest in quadrupeds, it should be considered in the case of such stooped senile individuals also.

SUMMARY

Studies of the thoracic form in 9,369 rural inhabitants of the Tōhoku area in Japan indicate certain developmental and retrogressive changes. The thorax of the newborn is rounder in transverse section but flattens increasingly until 4 years of age. In childhood, girls' chests are flatter than those of boys, but at the time of puberty, girls' chests suddenly become rounder and this is followed by a continuing tendency to greater

roundness with age. The chests of boys continually become flatter until the 25–29 year old period. These changes seem to be developmental. Despite differences in the thoracic index between occupational groups, all such adult groups show a gradual rounding of the thorax after the age of 30 and the change accelerates with age. Two reasons have been considered as a cause of the rounding of the thorax with age: decrease in stature, especially in sitting height, and senile curvature of the spine.

LITERATURE CITED

MARTIN, RUDOLF 1928 Lehrbuch der Anthropologie, 2nd ed., Gustav Fischer, Jena.

MOLLISON, TH. 1933 Phylogenie des Menschen. In Handbuch der Vererbungswissenschaft, edited by E. Baur and M. Hartmann, Bd. 3, Gebrüder Borntraeger, Berlin.

RELATIVE FAT PATTERNING: AN INDIVIDUAL CHARACTERISTIC *

BY STANLEY MARION GARN

Fels Research Institute, Antioch College
Yellow Springs, Ohio

INTRODUCTION

ANTHROPOLOGISTS, clinicians and biometricians familiar with the appearance of nude subjects or experienced in viewing body-build photographs have long been impressed by marked individual differences in subcutaneous fat patterning. Such differences, equally apparent when soft-tissue radiographs are reviewed, do not always escape notice even when subjects are fully dressed. A number of seemingly different fat "patterns" may be intuitively identified in even a small series of persons of like sex, constant age, and from a single racial group.

The medical and anthropological literatures are well stocked with descriptive terms for different subcutaneous fat patterns. Without attempting to list all of them, and excluding pathological conditions like *adiposis dolorosa,* and such unusual depositions of fat as characterize *steatopygia,* certain patterns have been listed repeatedly. They include an "upper extremity pattern," a "girdle fat pattern," and a "lower extremity pattern." Variants in thoraco-abdominal fattening have been described as "D" shaped and "d" shaped. Further we read of a "juvenile fat pattern," an "adult fat pattern" and a "Rubens fat pattern" (Martin '28, Sheldon *et al.,* '40, Angel '49, Škerlj *et al.,* '53).

* Section 3 of this paper is based on data obtained in the course of investigations carried out in collaboration with Dr. Josef Brožek at the Laboratory of Physiological Hygiene, University of Minnesota, and supported in part by funds provided under a contract between the regents of the University of Minnesota and the Department of the Army (contract No. DA 44-109-gm-1526). Permission to use the data was obtained from the Quartermaster Food and Container Institute for the Armed Forces and the Medical Nutrition Laboratory, United States Army. The views or conclusions contained in this report are those of the author. They are not to be construed as necessarily reflecting the views or endorsement of the Department of Defense.

While such designations are often devastatingly good, they lack quantification, depend much upon the judgment and preferred categories of the viewer, and may reflect in part configurations of the structures underlying the *tela adiposa* (*cf.* Škerlj *et al.*, '53, p. 593, Reynolds '51, p. 152). More important they may represent simple nutritional variants. As we have previously shown, the rate of fat deposition differs on different parts of the body so that (in absolute terms) the same individual may exhibit a different fat pattern when thin than he does when fat (Garn '54). Similarly, older and younger men may differ in the pattern formed by the absolute fat thicknesses not because of an age factor, but simply because the older man is fatter (Garn and Harper '55).

If, however, fat patterning is dealt with not in absolute terms, that is, the measured thicknesses themselves, but in terms of their z score transforms, pattern differences due solely to the amount of fat are eliminated. In this way it is possible to investigate interpersonal differences in fat patterning duly compensated for differences in the nutritive state. This technique of using z score profiles, which has had extensive use in psychometrics under the name of "profile analysis," seems highly applicable to the study of what we shall call *relative fat patterning*.[1]

It is the purpose of this study to investigate the extent of interindividual differences in "relative fat patterning." We shall indicate the degree to which individual males depart from the pattern expected for their degree of fatness. We shall demonstrate the existence of specific relative fat patterns in a series of adult males. And finally, we shall present evidence that individual relative fat patterns persist during weight loss, thus indicating that the relative fat pattern is an individual characteristic with some degree of permanence.

METHODS AND MATERIALS

The present study of relative fat patterning in the adult male is based on micrometer-caliper measurements of subcutaneous fat "shadows" as seen in soft-tissue X-rays of 94 adult males. The

[1] The z score is $(X - \bar{X})/\sigma$. In the present study the term "absolute fat pattern" as it applies to individuals refers to the pattern formed by the absolute or measured fat thicknesses, while the term "relative fat pattern" is used to designate the individual's z score profile. The absolute fat pattern may also be thought of as the "overt fat pattern," since it is this aspect of patterning that is perceived by the trained observer, while the relative fat pattern may be considered to be the "covert fat pattern," since it is ordinarily evident only after converting the absolute measurements into z scores.

roentgenographic techniques and the location of the fat thicknesses measured are as described by Garn and Saalberg ('53), Garn ('54) and Garn and Harper ('55).

Eighty-one of the subjects were adult male participants in the Fels Longitudinal Studies and were for the most part fathers of Fels children (Garn '54). Thirteen of the men were subjects in an investigation of controlled weight loss, conducted at the Laboratory of Physiological Hygiene, University of Minnesota (Brožek *et al.*, '55). These latter men were X-rayed by the author and Dr. Josef Brožek during a maintenance period and again at the end of a period of restricted caloric intake, employing the same roentgenographic procedures used on the larger series.

Each of the 9 measured fat thicknesses for each subject was converted into the corresponding z score or standard score value (see Guilford, '36) using the normative data published earlier (Garn '54).[2] In this way the 9 fat measurements, which differ markedly in thickness and absolute variability, were placed upon a comparable basis. Z scores, it may be recalled, have a mean of 0 and a standard deviation of *1* (Guilford '36, p. 86, Edwards, '47, p. 47); thus a hypothetical individual exactly 1 sigma above the mean in all 9 fat measurements would be represented by a z score value of $+1$ throughout. The appearance of such z score profiles is depicted in figure 1.

The extent of variability within an individual was measured by the statistic σz. This is simply the standard deviation of the individual's own z scores for subcutaneous fat. For the hypothetical individual described above σz would be 0 since all of his z scores were identical. Individuals more variable in relative fat patterning might have σz values of 0.5, 1.0 or even 1.5. σz therefore measures the extent of individual variability in subcutaneous fat patterning, corrected for the amount of fat present.

However, while σz provides a numerical measure of relative fat variability, it cannot of itself indicate where an individual departed from the general pattern. Accordingly, those individuals with above-median

[2] The use of z scores implies either that the distributions are normal or of the same form. This is not strictly true, in the present case, since χ^2 tests of the 36 combinations reveal a significant degree of heterogeneity of the distributions, with the distribution of hip fat somewhat different from 5 of the 8 others. However, this does not affect the general nature of the findings and especially the pattern comparisons. For a short discussion and further references see McNemar ('49, p. 38).

σz values were hand-sorted, and their z score values plotted on suitable coordinate paper, as shown in the lower part of figure 1. In this way it was possible to analyze, and in some cases categorize the diverse relative fat patterns making for high σz values. The "pattern index" technique developed by Lacey and Van Lehn ('52) was followed in part.

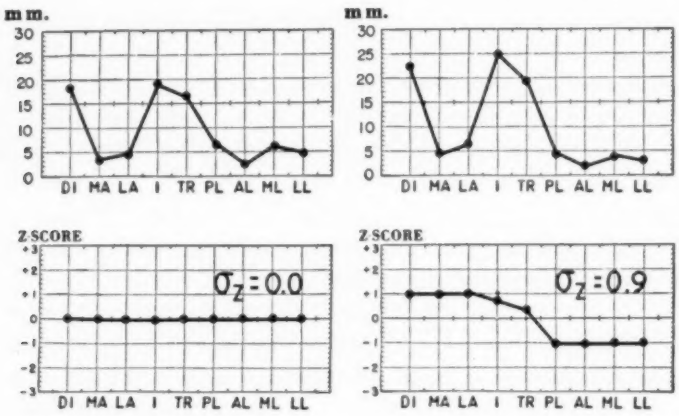

FIG. 1. Comparison of absolute fat patterns (above) and corresponding relative fat patterns (below). Left, hypothetical individual exactly at the mean in all 9 fat thicknesses. His mean z and fat variability score (σz) are both 0. Right, hypothetical individual with relatively thick arm fat and relatively thin leg fat. Though the difference between the two men is evident in the absolute fat patterns, it is striking in the relative patterns and in the fat variability scores.

The abbreviations: DI, MA, LA, I, TR, PL, AL, ML, and LL refer to deltoid insertion fat, medial arm fat, lateral arm fat, iliac fat, trochanteric fat, posterior leg fat, anterior leg fat, medial leg fat and lateral leg fat.

Finally the permanence of such z score fat patterns, or more precisely their resistance to change during weight loss, was investigated in the Minnesota group. Here the z score profile patterns before and after weight loss for each individual were plotted and compared. The degree of similarity or fit between the "before" and "after" patterns was expressed numerically as σdz, the standard deviation of the z score differences. Like σz, σdz also has a theoretical lower limit of 0, and no upper limit. A σdz of 0 would be attained if an individual in losing fat lost the same amount (in z score units) in all 9 fat thicknesses

measured. Thus σdz measures the degree of *similarity* between z score profile patterns, and may be called the *profile similarity index.*

These procedures which deal with z scores, rather than absolute fat thicknesses, are made necessary by the facts that the different fat thicknesses measured differ considerably in their mean values, and that the rate of fat deposition or withdrawal is non-uniform over different parts of the body.

FINDINGS

1. Relative fat variability scores (σz)

To begin with, the standard deviation of individual z scores for subcutaneous fat (σz) was calculated for each of the 81 males in the basic series. As shown in table 1, the value of σz ranges from a low of 0.26 to a high of 1.45. The lowest value observed approaches the

TABLE 1

Distribution of Relative Fat Variability Score (σz) *for 81 Adult Men*

RANGE	NO.	PERCENT	CUMULATIVE PERCENTAGE
0.00-0.19	0	0	0
0.20-0.39	11	14	14
0.40-0.59	37	46	60
0.60-0.79	18	22	82
0.80-0.99	10	12	94
1.00-1.19	2	2	96
1.20-1.39	1	1	97
1.40-1.59	2	2	99
1.60-1.79	0	0	99

The mean is 0.61 and the median is 0.56.

theoretical minimum of 0.0, and indicates a high degree of z score uniformity, and an almost total absence of relative fat patterning. The high values, on the other hand, are indicative of considerable internal variability, disproportion or " dysplasia " in subcutaneous fat deposition. Since σz is a new statistic in this work, it is of interest to note that the mean was 0.61, and the distribution considerably skewed, with a Pearsonian coefficient of skewness (Sk) of 0.75. While such skewness is to be expected where fat is concerned, it indicates that the median rather than the mean σz should be employed as a measure of central tendency. Accordingly, in selecting individuals with high σz values, for the second

part of the study, selection was based on above-median rather than above-average σz scores.

Though, anthroposcopically, older men seem to be more variable in absolute fat patterning, there was no evidence that age alone influenced the relative fat variability scores. The mean σz for 40 younger men did not differ significantly from the mean for 41 older men. As a confirmation the product moment r between age and σz failed to attain significance at the 5% level. In like fashion, thin and fat men did not

TABLE 2

Fat Variability Scores (σz) for Older and Younger and Fatter and Thinner Men

GROUP	NO.	MEAN AGE	σz RANGE	MEAN σz	S. D.
Total group	81	39.8	0.26-1.45	0.61	0.24
Younger men	38	30.4	0.26-1.08	0.62	0.21
Older men	43	48.1	0.32-1.45	0.61 [3]	0.26
Leaner men [1]	41	39.5	0.28-1.45	0.57	0.25
Fatter men [2]	37	40.8	0.32-1.43	0.66 [3]	0.25

[1] Trochanteric fat below 16.6 mm.
[2] Trochanteric fat over 16.6 mm.
[3] Differences between means not significant at 5% level.

differ significantly in mean σz (table 2). However, a low, but positive and significant correlation of 0.32 between trochanteric fat and σz suggested at least that fatter, or fatter-hipped, men may be slightly more variable in relative fat patterning. Under these circumstances, some degree of caution may be required, before saying that the very fat are more "dysplastic" than the very thin.

2. Relative fat patterns

Since a given value of σz merely expresses the degree of variability in the individual's z scores for fat, without explaining the source of the variability, it was of interest to explore the various relative fat patterns in order to determine whether there were "upper extremity," "lower extremity" and "girdle fat patterns," etc. Therefore the z score values for subcutaneous fat were plotted for all men above the median σz, and the z score profile patterns were examined further.

Of the 40 men so plotted, it was not possible to describe or "name" the pattern anatomically in 23 cases. That is, these 23 men high in σz were by inspection high or low in individual fat thicknesses (medial arm,

posterior leg, etc.) or showed highly localized highs and lows. In 17 men, however, the patterns involved anatomical regions, with deviations up to two z scores from their own mean z and ranges up to 4 z scores.

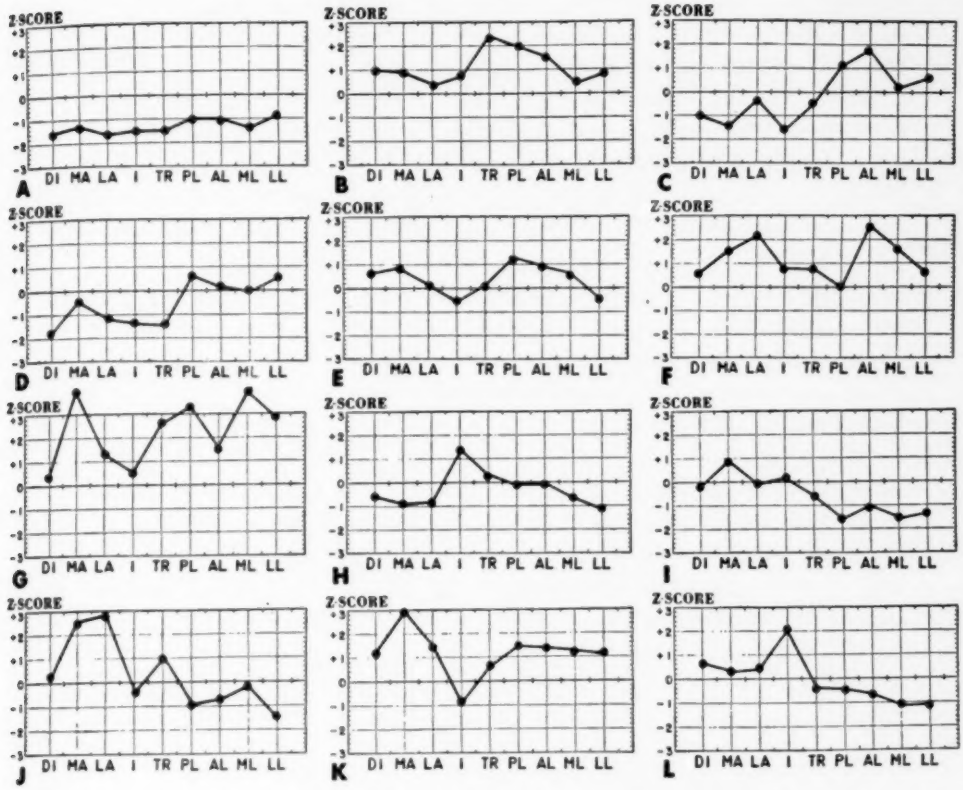

FIG. 2. Twelve representative z score profile patterns for fat. A, near-absence of relative fat patterning. B, lower segment pattern. C and D, leg fat patterns. E and F, arm and leg patterns. G, example of a pattern not easily categorized. H, iliac fat pattern. I, arm fat pattern. J, arm-trochanter fat pattern. K, low hip-fat pattern (-2σ below the individual's mean). L, high hip pattern (1.9σ above individual's own mean).

The abbreviation DI, MA, LA, etc. refer to specific fat thicknesses as in figure 1.

Such cases where lower leg fat or arm fat responded as a whole were simpler to name.

In 5 men there was a clearly-marked leg fat pattern characterized by elevated z score values for the entire leg region. Examples are shown in figure 2, c and d, wherein it is evident that a leg fat pattern

(averaging two or more z scores above the individual's average z) may be found both in men who are thin and those others who are fat. Three additional men showed a combined arm and leg pattern, giving their z score profiles a notably sinusoidal appearance. One man, out of the 40, evidenced a combined pelvis and leg pattern, a combination rarely to be expected on a chance basis alone.

In all, 3 men exhibited the kind of concomitant trochanteric-iliac fat excess, that could be called a "girdle fat pattern." However, it should be noted that the z scores correct for the usual tendency to deposit fat differentially in this region, so that these 3 men are exceptionally "hippy" specimens. The individual pictured in figure 2, h, is average in his total fat thickness but markedly deviant in iliac fat.

Other nameable patterns include an arm-hip pattern (figure 2, i), a true arm pattern (figure 2, j), and a pattern involving low iliac fat (figure 2, k). The individuals shown in figure 2 k and l are quite exceptional in that their iliac fat is nearly 2 standard scores out of line from their *own* z score mean.

Therefore it is evident that there are a number of relative fat patterns contributing to an above average σz, and characteristic of individuals. While some of these relative fat patterns can be described by terms previously used (lower-extremity, upper-extremity etc.) over half of them are not easily described in words. Under these circumstances the z score profiles, or better, t score profiles, may prove more useful in their original form than when further categorized. To put it differently, there are more, and more complicated, relative fat patterns than we can conveniently force into categories.

3. Permanence of the relative fat pattern

While it is evident that there are marked interpersonal differences in relative fat patterning, one may well ask whether any given relative fat pattern possesses permanence, and persists despite changes in nutritional status. Fortunately data on the 13 young men from the weight loss series allow us to answer this question with some degree of assurance.

As shown in figure 3, 1 through 13, the individual z score profile patterns after weight loss closely resemble the initial patterns and in several cases are practically identical (figure 3, 1, 2, 3, and 12). In none of the 13 cases did food restriction cause a reversal of pattern or even a marked change. Interestingly, the tendency for the individual relative fat patterns to hold, despite generalized loss of fat, is as marked in those men who were initally obese as among those who were lean even before food restriction.

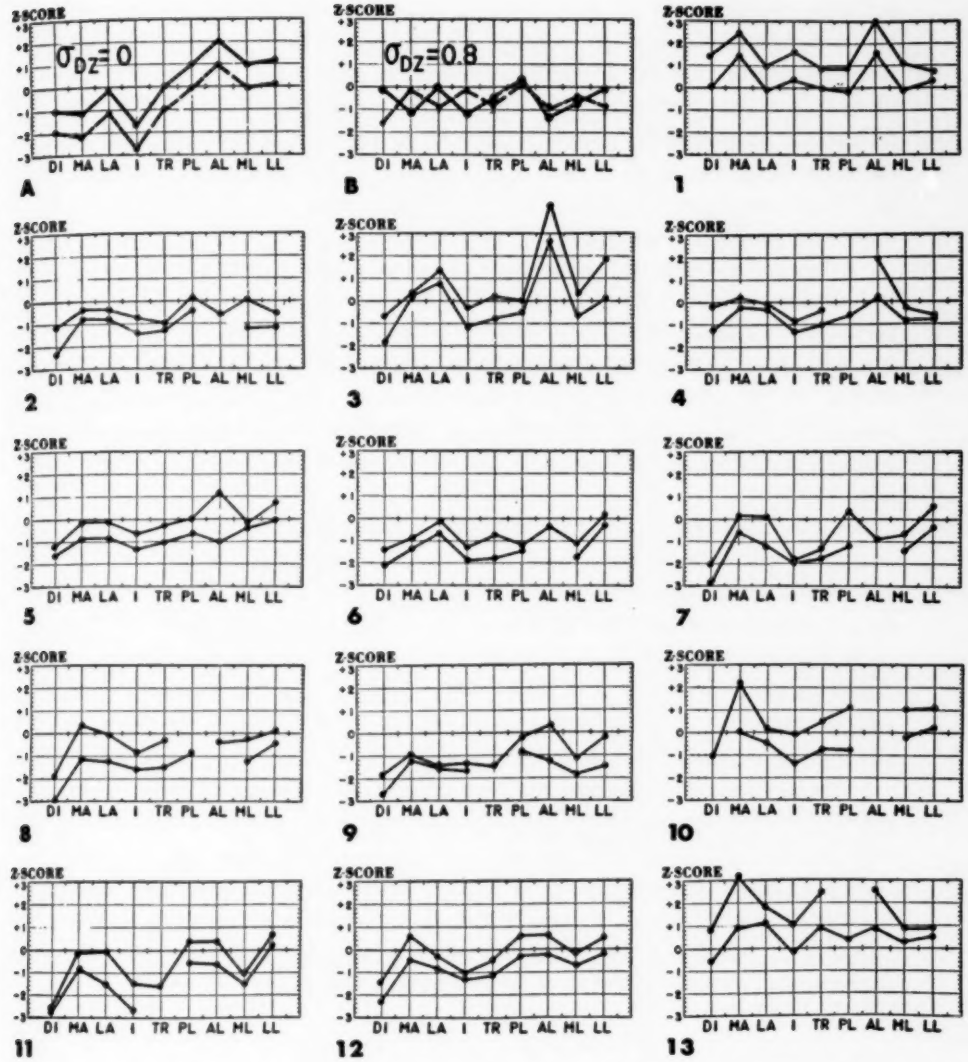

FIG. 3. 1-13; Relative fat patterns of 13 young men before and after a period of food restriction. Despite generalized fat loss, as shown by the lower level of each "after" pattern, individual pattern profiles maintained their identity.

A and B; hypothetical patterns illustrating (A) perfect pattern congruence and (B) pattern dissimilarity with several reversals.

However, while the "before" and "after" z score profiles provide striking visual proof of the resistance of the relative fat pattern to changed nutritional status, the plotted patterns do not afford a quantitative measure of pattern similarity. Therefore, quantification was accomplished by computing σdz which is the standard deviation of the differences expressed in z score units. As pointed out previously, and as illustrated in figure 3, A and B, if two z score profile patterns are identical, the value of σdz will be 0. σdz, therefore, measures the degree of *profile similarity.*

As shown in table 3, the observed values of σdz (the profile similarity index) range from 0.20 to 0.50 for the 13 men in the weight reduction series, with a mean of 0.39 $\pm$ .03 for the group as a whole. In contrast the values of σdz for 13 randomly selected pairs of men were much higher, ranging from 0.50 to 1.36 and the mean σdz for these randomly selected pairs was 0.77 $\pm$ .07. Thus an individual continues to resemble himself in the relative fat pattern, despite considerable and rapid weight loss.

Since all of the men in the weight-loss series had lost fat during the course of the experiment, there were possible objections to the use of random pairs for comparison, even though the actual pairings were made by a disinterested member of the Fels Institute staff. Accordingly, a second set of pairs was set up, matched on a weight basis to the starvation series men before and after food restriction. Here, the values of σdz ranged from 0.41 to 1.47, with a mean of 0.67 $\pm$ .07. These results were substantially similar to those mentioned above.

These particular findings may be summarized by saying that individual relative fat patterns hold their shape during weight reduction even though the measured fat thicknesses (*i. e.,* the absolute fat patterns) change. The use of σdz makes it possible to demonstrate that the relative fat patterns of an individual under two different conditions of nutriture are more alike than the relative fat patterns of two different individuals, whether selected on a weight basis or not.

DISCUSSION

In this communication we have returned to a consideration of the individual, the prime unit of constitutional research, though limiting ourselves to the study of individual differences in one tissue layer (the *tela adiposa*), one sex and age group (adult males), and to purely objective measurements. We have shown that it is possible to go beyond

TABLE 3

Profile similarity indices for weight-reduction series, random pairs and weight matched pairs

Subject or Pair Number	I WEIGHT-LOSS SERIES Weight loss Kg	*σdz*	II RANDOMLY-MATCHED PAIRS Weight difference Kg	*σdz*	III WEIGHT-MATCHED PAIRS [1] Weight difference Kg	*σdz*
1.	14.4	0.34	— 12.1	0.69	7.6	0.79
2.	9.4	0.42	5.5	0.62	10.0	0.67
3.	10.2	0.48	— 13.2	1.36	10.4	0.62
4.	8.9	0.48	— 0.9	0.63	8.5	0.76
5.	9.8	0.28	— 6.6	0.73	8.8	0.68
6.	8.3	0.20	— 13.3	0.57	7.9	1.47
7.	8.9	0.43	— 26.8	0.55	9.5	0.47
8.	10.0	0.30	6.9	0.76	10.0	0.66
9.	6.9	0.48	— 7.9	0.50	6.7	0.81
10.	9.4	0.53	3.9	0.61	9.5	0.47
11.	5.0	0.39	0.8	0.93	8.8	0.41
12.	11.3	0.24	—1.4	1.01	9.8	0.65
13.	6.9	0.50	26.0	1.09	15.0	0.51
	9.2 ± .64	0.39 ± .03	— 3.01 ± 3.51	0.77 ± .07	9.4 ± .54	0.69 ± .07

[1] Matched to initial and terminal weights of group I.

the absolute fat patterns, which are highly responsive to changes in caloric intake, to the relative fat patterns, which appear to be of a more permanent nature.

The measure σz, the standard deviation of the individual's own z scores for fat, provides a direct measurement of relative fat variability. σz, which has a lower limit of 0, and no upper limit, measures the extent to which the individual departs from a straightline z score profile. Since z score units are used, and variability is measured about the individual's own mean z, age does not enter into the picture, and for the most part differences in the degree of obesity may be ignored. A male with a σz greater than 0.61, or better 0.90, may be considered dysplastic (in regard to distribution of fat) whether fat or lean.

While a high σz indicates that an individual does have an unusual fat pattern, it cannot of itself indicate why or where. Here, actually plotting the z score profiles for fat becomes a necessity. As shown by the analysis of individual z score profiles, there are a number of distinct and individually characteristic relative fat profiles. Some of them can be categorized as "arm-fat," "leg-fat," "arm and leg fat," etc. On the other hand, there are a larger number, quite distinct in the profiles, that cannot be categorized, or categorized in simple language. And, since our interests extend well beyond the level of simple description there is no need to do so.

The probability that a given relative fat pattern is an individual characteristic, possessing some degree of permanence through time, is heightened by the findings in the weight reduction group. Here all of the men lost weight (up to 14 kg) and all lost fat in all 9 fat layers measured. Consequently the z score patterns were all displaced downward at the end of the weight loss period. Yet the individual patterns tended to "hold." The extent of similarity between the individual "before" and "after" patterns was striking, and there was little difficulty in recognizing each individual from his post-starvation z score pattern. However, the permanence of such patterns has not yet been demonstrated over longer time periods, or under different nutritional regimens.

Still, it is important to point out the utility of the relative fat pattern concept. As is well known, absolute fat patterns undergo qualitative changes during adolescence and less striking quantitative changes thereafter. Under these circumstances, permanence of type cannot be sought either in the absolute fat thicknesses, or in the patterns

formed by the absolute fat thicknesses. Relative fat patterning seems to afford the only approach both to the problem of permanence of type and the problem of genetic determinants, as far as the *tela adiposa* is concerned.

Previous workers have demonstrated a general tendency for children to hold position, either on a rank basis or a percentile basis, in the thickness of specific fat layers during the growing period (Stuart *et al.*, '40, Stuart and Sobel '46, Reynolds '51). While such constancy suggests a genetic determinant, it is equally likely that children maintain their eating habits over the years: a juvenile trencherman continuing to be a trencherman and a child with a bird-like appetite continuing his avian behavior. Reynolds ('51) further demonstrated the existence of sibling similarities in the thickness of the *tela adiposa*. Again, while the possibility of a genetic determinant cannot be ignored, the fact that siblings share a common diet should not be discounted either. Under these circumstances the relative fat pattern may prove of maximum use both in studying constancy of type and in searching for genetically-determined variants in the *tela subcutanea*.

Finally, it should be pointed out that the relative fat pattern does not replace the absolute fat pattern. The absolute fat pattern comes closest to what we perceive. The absolute fat pattern is affected by age, sex, and nutritional status, and it is more than likely that the absolute fat patterns of different races differ to an important degree. The relative fat pattern, on the other hand, is much more of an individual characteristic and, in adulthood at least, is seemingly much more resistant to change. Both aspects of fat patterning are important adjuncts to research on the subcutaneous fat. Absolute fat patterns obviously relate to nutritional studies. Relative fat patterns, on the other hand, afford an opportunity for the investigation of the individual and are far more likely to find their origin in the germ plasm than in the "groaning board."

SUMMARY

1. Individual differences in *relative fat patterning* were investigated after converting absolute measurements into their corresponding z transforms. In this way patterned differences due solely to the amount of subcutaneous fat were eliminated.

2. The measure σz (the standard deviation of the individual's own z scores for fat) was used to indicate the extent of *relative fat variability*.

In 81 adult males σz ranged from 0.29 to 1.45 with a mean of 0.61. It was suggested that σz could serve as an objective measure of fat "dysplasia."

3. As tested in various ways, σz was not affected by age and only slightly, if at all, by the total amount of fat.

4. Individual relative fat patterns were investigated in the z score plots of 40 men with above-median σz values. A number of characteristic relative fat patterns were found, some of them resembling patterns described in the literature. In 17 men there were distinct and nameable relative fat patterns. However, in 23 cases, the relative fat patterns were not susceptible to simple naming.

5. The possibility that such relative fat patterns are permanent and resistant to changes in nutritional status was investigated by comparing the z score profile patterns of 13 young men during a period of food restriction.

6. Visually, the tendency of an individual to maintain his z score profile pattern was striking. Using a quantitative measure of pattern similarity (σdz) this impression was confirmed.

7. It was observed that the relative fat pattern is an individual characteristic, having some permanence and resistance to change under nutritional stress, and that the relative fat pattern (in contrast to the absolute fat pattern) may be genetically determined.

ACKNOWLEDGMENTS

The author expresses his indebtedness to Dr. John I. Lacey, whose "profile pattern" technique has been adopted in the present study and whose statistical suggestions proved most helpful throughout. Mr. Irvin Emanuel should be given credit for the use of σz and the suggestion that σz scores for fat serve as a measure of fat "dysplasia." The helpful cooperation of Dr. Josef Brožek, Mrs. Nedra Foster, and the staff of the Laboratory of Physiological Hygiene of the University of Minnesota is gratefully acknowledged.

Further, the technical assistance of Richard Young and Georgia Rasmussen Greaney and the aid of Ruth V. Harper during the conception and realization of the study should be mentioned. Lois A. Conklin aided in the manuscript preparation.

LITERATURE CITED

ANGEL, J. L. 1949 Constitution in female obesity. Am. J. Phys. Anthrop., n. s. 7: 433-471.

EDWARDS, A. L. 1947 Statistical Analysis. Rinehart and Co., New York.

BROŽEK, J. *et al.* 1955 Changes in body composition during maintenance on survival rations (to be published).

GARN, S. M. 1954 Fat patterning and fat intercorrelations in the adult male. Human Biol., *26*: 59-69.

GARN, S. M. AND R. V. HARPER 1955 Fat accumulation and weight gain in the adult male. Human Biol., 27: 39-49.

GARN, S. M. AND J. H. SAALBERG 1953 Sex and age differences in the composition of the human leg. Human Biol., *25*: 144-153.

GUILFORD, J. P. 1936 Psychometric Methods. McGraw-Hill Book Co., New York.

LACEY, J. I. AND R. VAN LEHN 1952 Differential emphasis in somatic response to stress. Psychosom. Med., *14*: 77-81.

MARTIN, R. 1928 Lehrbuch der Anthropologie, 2nd. ed., G. Fischer, Jena.

MCNEMAR, Q. 1949 Psychological Statistics. John Wiley, New York.

REYNOLDS, E. L. 1951 The distribution of subcutaneous fat in childhood and adolescence. Monographs of the Society for Research in Child Development. 15 (2): Serial No. 50.

SHELDON, W. H., S. S. STEVENS AND W. B. TUCKER 1940 The Varieties of Human Physique. Harper and Brothers, New York.

ŠKERLJ, B., J. BROŽEK AND E. E. HUNT, JR. 1953 Subcutaneous fat and age changes in body build and body form in women. Am. J. Phys. Anthrop. n. s., *11*: 577-600.

STUART, H. C., P. HILL AND C. SHAW 1940 The growth of bone, muscle and overlying tissues as revealed by studies of roentgenograms of the leg area. Monographs of the Society for Research in Child Development. 5 (3): Serial No. 26.

STUART, H. C. AND E. H. SOBEL 1946 The thickness of the skin and subcutaneous tissue by age and sex in childhood. J. Pediat., *28*: 637-647.

A MARKOV PROCESS MODEL FOR FOLLOW-UP STUDIES

BY SAMUEL ZAHL

Department of Biostatistics,
Harvard School of Public Health

1. INTRODUCTION

THE purpose of this paper is to introduce the general, time stationary, continuous, Markov process with a finite number of states, as a model for the follow-up study of a disease. A special case of such a process was introduced by Fix and Neyman ('51) as a model for the study of cancer or other disease.

In models used up to the present time cure rates or death rates have generally been used to measure the effectiveness of a treatment. These rates may be too coarse in that they overlook other important characteristics of the disease or treatment. They may also be not quite relevant in that the disease may cause death only indirectly and only after many years, or else the treatment does not cure in the sense of the patient becoming free of at least the gross symptoms, but only retards the progress of the disease. This is the situation with many, if not most, of the degenerative diseases. Under these circumstances, a medical scientist will most likely incorporate intuitive and non-quantitative measures into his appraisal of the treatment.

To introduce other quantitative measures of a treatment, it is necessary to consider more elaborate models of the temporal progress of the disease than have hitherto been used. One such elaboration, where it is appropriate to do so, is to define several states that a patient may be in. For example, in studying arthritis we might define the following states: 1, active arthritis; 2, quiescent; 3, death from arthritis; 4, lost from observation or death from other causes. This might be enlarged by replacing state 1 with several states representing a range of activity from active to quiescent. The exact number of states would depend on

the purposes of the investigation and the ability of the investigator to discriminate between the states. Given such a set of states the patient who enters the study is regarded as traversing a path: he enters the study in some initial state and from time to time transfers from one state to another and eventually leaves the study in some final state. It is assumed that the paths are determined by a set of stochastic or probabilistic laws, *i. e.*, the laws specify the probability of any particular path, though not the path itself. With the aid of such a model the disease process is studied as a stochastic process. The new measures which it is hoped are useful are derived from certain characteristics of the stochastic process. One such characteristic is the probability that a patient who enters the study classified in one particular state will, after T years, leave the study in some other, or the same, particular state. Such a T year transition probability, as it is called, is a generalization of the more familiar T year survival rate. If there are k states, then there are k^2 such T year transition probabilities. As a second characteristic we have, for all patients who enter the study in one particular state and leave the study after T years, the average time spent in each of the states. As an illustration of this second characteristic, suppose a drug is being tested for its power of inhibiting the activity of arthritis. Then the average length of time which a person spends in a less active state of arthritis, who originally received treatment in a more active state, might be a measure of the effectiveness of the drug.

The chief statistical problem is to estimate the set of parameters which defines the Markov process, by means of a sample of paths. The sampling scheme considered in this paper is one which consists of periodic observations on patients, that is, if a patient enters the study at time t_0, he is observed at times $t_0 + t, t_0 + 2t, t_0 + 3t$, etc., and only the information concerning the state of the patient at these times is used to estimate the parameters of the process.

In section 2 a brief description of the Markov process used as a model is given. In section 3 certain useful characteristics of the process are defined. These sections are included for the sake of completeness only. A more thorough treatment of the mathematical aspects of the model is given, for example, by Doob ('53), or Feller ('50) both of whom give extensive bibliographies. In section 4, the estimates of the parameters (called transition rates) of the model are given. Estimates of the characteristics are then derived from the estimates of the transition rates. Section 5 treats of the large sample distribution of the estimates, and the

next two sections, 6 and 7, of tests of hypotheses and tests of goodness of fit of the model, respectively. In section 8 the model is generalized to two contiguous study periods. In section 9 an example is given.

The terminology used in this paper refers to disease processes. As far as the theory is concerned, this is only a matter of convenience and is not meant to imply any limitation to the study of disease alone.

2. A BRIEF DESCRIPTION OF THE MODEL

We assume that during the course of a follow-up study a patient is at every instant of time in one and only one of k states which we label $S_1, S_2, \cdots, S_k$. The follow-up study is of length T, which means that each patient is under observation for a period of time T following his entry into the study. The actual starting times of the patients are ignored and the entry times all set equal to zero. The total period of the study is therefore from $t=0$ to $t=T$, and the fact that a patient is assumed to be in one of the k states at each instant of this interval of time makes the process, by definition, a continuous one.

The phrase, under observation, in the context of this paper, does not mean strictly physical observation; it means rather, counted as being in the study and capable of being classified in one of the k states at any instant of time, while in the study. By this definition, for example, a patient who dies during the study remains formally under observation until his time T has elapsed, if death is one of the states. We remark also that the times $t=0$ to $t=T$ need not, in practice, represent the actual length of the follow-up study. These two times may, for example, represent years 0 to 5, or 6 to 10, or any interval of a 20 year follow-up study.

We are concerned with certain events called transitions. A transition is an event of the following form: a patient at time t_1 is in state S_i and at time $t_2 > t_1$ is in state S_j, where i and j may be equal or unequal. Nothing is said about the state of the patient at any other time. Associated with these events are probabilities. If a patient is in state S_i at any time t_0 we assume that there exists a conditional probability $p_{ij}(t)$ that at a later time $t_0 + t$ the patient will be in state S_j, for $i, j = 1, \cdots, k$ and $0 \leq t_0 \leq t_0 + t \leq T$. The probabilities $p_{ij}(t)$ are called transition probabilities. We have defined the transition probability to depend only on the difference $(t_0 + t) - t_0 = t$ and not on the actual time points t_0 and $t_0 + t$; by definition this is the property of stationarity. When making statements about $p_{ij}(t)$ for all values of t

in the interval $[0, T]$ we will also refer to $p_{ij}(t)$ as a transition probability, in order not to complicate the terminology too much, though strictly speaking, $p_{ij}(t)$ should then be called a transition probability function.

The additional assumptions that characterize our model are as follows. When $t = 0$,

$$p_{ij}(0) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}. \tag{2.1}$$

The transition probabilities are also continuous with respect to t at $t = 0$, *i.e.*,

$$\lim_{t \to 0} p_{ij}(t) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}. \tag{2.2}$$

The last assumption is that our process has the Markov property. Loosely speaking this means that the conditional probability of the transition from S_i to S_j in time t does not depend on the history of the patient previous to the time he was in state S_i. More precisely, if $p_i(0)$, $i = 1, \cdots, k$ is the probability of being in state S_i at $t = 0$, and $0 < t_1 < t_2 < t_3 < \cdots < t_n < T$ are any finite set of time points, then the probability that a patient is in state S_{i_0} at $t = 0$, S_{i_1} at $t = t_1, \cdots, S_{i_n}$ at $t = t_n$ is

$$p_{i_0}(0) p_{i_0 i_1}(t_1) p_{i_1 i_2}(t_2 - t_1) \cdots p_{i_{n-1} i_n}(t_n - t_{n-1}).$$

The set of transition probabilities $p_{ij}(t)$ are conveniently assembled into a transition matrix $P(t) = \{p_{ij}(t)\}$. We observe that when $t = 0$, $P(0)$ is the identity matrix. Matrix terminology is very convenient for our presentation and it will be used extensively.

It can be shown (Doob '53, pp. 238-243) that there exists a unique matrix Q such that for $0 \leq t \leq T$,

$$P(t) = I + Qt + Q^2 \frac{t^2}{2!} + Q^3 \frac{t^3}{3!} + \cdots = e^{Qt} \tag{2.3}$$

symbolically. Q does not depend on t; its elements are the instantaneous transition rates, as may be seen from the relation

$$\lim_{t \to 0} \frac{1}{t}(P(t) - I) = Q.$$

The elements of Q will therefore be called transition rates and Q the transition rate matrix, or simply rate matrix. Q together with the

initial states of the patients determines the process, so the primary estimation problem is to estimate Q.

The transition rates have the following properties:

$$q_{ij}\begin{cases}\geq 0 & i\neq j\\ \leq 0 & i=j\end{cases} \tag{2.4}$$

$$\sum_{j=1}^{k} q_{ij}=0 \qquad i=1,\cdots,k. \tag{2.5}$$

Any square matrix Q with the properties (2.4) and (2.5) is the rate matrix of some process (of the type we are concerned with). This is a useful property enabling us to manufacture processes at will.

In conclusion, the most important conditions which the disease process must satisfy, at least approximately, in order for the model to apply, are the following:

1. During the period $[0,T]$ of the study, the patient is at any time in just one of a finite number of clearly defined states, $S_1,\cdots,S_k$.

2. The probability that a patient in state S_i $(i=1,\cdots,k)$ at any time during the study, is subsequently in state S_j $(j=1,\cdots,k)$ at a later time in the study, depends on the two states in question but not on the previous history of the patient.

3. The probability that a patient in state S_i at time t_1 is in state S_j at a later time t_2 depends on the difference t_2-t_1, but not on the initial time t_1.

3. CHARACTERISTICS OF THE DISEASE PROCESS

Starting with a given rate matrix Q, which may be an estimate or else theoretically given, we can derive certain characteristics of the process.

a. From (2.3) we can determine the transition probabilities for any time interval, in particular for the total period $[0,T]$, by letting $t=T$.

b. It may sometimes be desirable to find the transition probabilities corresponding, not to the given Q, but to a rate matrix obtained from Q by changing the values of some of the elements of Q, while at the same time satisfying the restrictions (2.4) and (2.5). As an important particular case, suppose the state S_δ is the state of being lost from observation. The rates $q_{i\delta}$ $(i=1,\cdots,k;i\neq\delta)$, representing rates of transfer from the other states to S_δ, will affect the transition probabilities in a way which, we usually assume, has nothing to do with the disease.

By substituting 0 for these q_{ij}'s and altering the diagonal elements of Q to satisfy (2.5), a new transition matrix can be calculated for the hypothetical situation in which no patients are lost from observation. In general, if the transition rates $q_{\alpha\beta}, q_{\gamma\delta}, \cdots, q_{\mu\nu}$ have been suppressed in this way, then the altered rate matrix will be called the net rate matrix and written $Q_{\alpha\beta|\gamma\delta|\cdots|\mu\nu}$, and its elements called net transition rates. The corresponding transition matrix will be called the net transition matrix and written $P_{\alpha\beta|\gamma\delta|\cdots|\mu\nu}(t)$, and its elements called net transition probabilities.

c. From a given rate matrix another characteristic that can be derived is: for each pair of states S_i, S_j, the expected time spent in state S_j during the period $[0, T_0]$, $0 \leq T_0 \leq T$, for a patient who is initially in state S_i. This expectation will be noted by $e_{ij}(T_0)$; it is given by

$$e_{ij}(T_0) = \int_0^{T_0} p_{ij}(t)\,dt$$

where $p_{ij}(t)$ is determined by Q from (2.3). The integral is the limit as $N \to \infty$ and $\| t_n - t_{n-1} \| \to 0$ of the approximating sum

$$\sum_{n=1}^{N} p_{ij}(t_{n-1})\, p_{ij}(t_n - t_{n-1}) \cdot (t_n - t_{n-1})$$

where $0 = t_0 < t_1 < \cdots < t_n = T_0$.

The matrix of expectations

$$e(T_0) = \{e_{ij}(T_0)\} \qquad i, j = 1, \cdots, k$$

may be computed using (2.3). We have

$$\begin{aligned} e(T_0) &= \int_0^{T_0} P(t)\,dt \\ &= \int_0^{T_0} I\,dt + Q\int_0^{T_0} t\,dt + Q^2 \int_0^{T_0} \frac{t^2}{2!}\,dt + \cdots \\ &= IT_0 + Q\frac{T_0^2}{2!} + Q^2\frac{T_0^3}{3!} + \cdots . \end{aligned} \tag{3.1}$$

The matrix $e(T_0)$ will be called the residence matrix and its elements, residencies. If $Q = Q_{\alpha\beta|\cdots|\mu\nu}$ is substituted in (3.1), then the resulting residence matrix will be called the net residence matrix and written $e_{\alpha\beta|\cdots|\mu\nu}(T_0)$. Its elements will be called net residencies.

It is clear from the above that the transition and residence matrices

depend entirely on Q. It is convenient therefore to estimate Q first and then to estimate the transition and residence matrices using the relations (2.3) and (3.1).

4. ESTIMATION IN THE MODEL

The estimate of Q is based on a series expansion of Q in terms of $P(t)$. Writing $P^*(t) = P(t) - I$, (2.3) becomes

$$I + P^*(t) = e^{Qt}. \tag{4.1}$$

Proceeding heuristically we then write

$$Qt = \log(I + P^*(t)) = P^*(t) - \frac{(P^*(t))^2}{2} + \frac{(P^*(t))^3}{3} - \cdots. \tag{4.2}$$

In appendix 1 it is proved that (4.1) and (4.2) are equivalent provided the series in (4.2) converges elementwise. A necessary and sufficient condition that the series converges is that the roots of $P^*(t)$ are all less than one in absolute value (Wedderburn, '34). This condition is not very convenient to apply in practice as it may involve considerable computation. For our purposes the following condition is more useful. If

$$p_{ii}(t) > \tfrac{1}{2} \text{ for } i = 1, \cdots, k \tag{4.4}$$

then the series converges. The proof is given in appendix 2.

Let us suppose now that at periodic intervals t' of time, starting from their respective times of entry into the study, patients are classified into one of the states $S_1, \cdots, S_k$. If a particular patient is classified s times in this way, his observed path will consist of the sequence of states $S_{i_0}, S_{i_1}, \cdots, S_{i_s}$, where S_{i_0} is the initial state, S_{i_1} the state after time t' has elapsed, etc. The likelihood of this path is

$$\pi_{i_0} p_{i_0 i_1}(t') p_{i_1 i_2}(t') \cdots p_{i_{s-1} i_s}(t')$$

where π_{i_0} is the probability of the patient entering the study in state S_{i_0}, while the remaining probabilities correspond to the succeeding transitions. Assuming that patients behave independently, the likelihood of the observed paths of n patients, after rearranging, is therefore

$$L = \prod_{i=1}^{k} \pi_i^{n_i} \prod_{i,j=1}^{k} p_{ij}(t')^{n_{ij}} \tag{4.5}$$

where n_i is the number of patients entering in state S_i, n_{ij} is the total

number of transitions $S_i \to S_j$ and $n = n_1 + \cdots + n_k$. In order that the likelihood not vanish we define $\pi_i^{n_i} = 1$ if $n_i = 0$ and similarly we define $p_{ij}(t')^{n_{ij}} = 1$ if $n_{ij} = 0$. A straightforward computation yields the maximum likelihood estimate of $p_{ij}(t')$:

$$\hat{p}_{ij}(t') = \frac{n_{ij}}{n_{i1} + \cdots + n_{ik}} \qquad i, j = 1, \cdots, k. \tag{4.6}$$

We observe that the π_i's do not enter, so that the estimating procedure is independent of the particular scheme of entry of the patients. We also observe that in order for $\hat{p}_{ij}(t')$ to be defined, the denominator in (4.6) must not vanish. The interpretation of this fact is that somewhere in the course of the follow-up study patients must be observed to be in each of the designated states.

In estimating Q we consider first the case that there are no *a priori* restrictions on the elements of Q, and second that there are restrictions of the form $q_{ij} = 0$ for certain of the transition rates. These restrictions would occur when it is known that the corresponding transitions $S_i \to S_j$ cannot occur.

Case 1. No *a priori* restrictions on the rates. Letting

$$P(t') = \{\hat{p}_{ij}(t')\} \quad \text{and} \quad \hat{P}^*(t') = \hat{P}(t') - I,$$

the estimate of Q is taken to be

$$Q = \frac{1}{t'}\left[\hat{P}^*(t') - \tfrac{1}{2}(\hat{P}^*(t'))^2 + \tfrac{1}{3}(\hat{P}^*(t'))^3 - \cdots\right] \tag{4.7}$$

provided the series converges. If the series converges, $\hat{Q}$ is the unique maximum likelihood estimates of Q. It is not certain, however, that the series will converge even if the roots of $P^*(t')$ are less than one in absolute value, since $\hat{P}^*(t')$ is a sample matrix and some of its roots may, by chance, exceed one. However, the probability of the series not converging can be made arbitrarily small by choosing t' sufficiently small. Condition (2.2) states that the parameters $p_{ii}(t')$ converge to 1 as $t' \to 0$, and the $\hat{p}_{ii}(t')$'s, being maximum likelihood estimates of the $p_{ii}(t')$'s, converge in probability to the latter. Hence, a practical solution is to choose t' so that the $p_{ii}(t')$'s are (sufficiently) greater that $\frac{1}{2}$. This will make the probability that any of the $\hat{p}_{ii}(t')$'s exceed $\frac{1}{2}$ (and the series possibly not converging thereby) small. Some *a priori* information as to the magnitudes of the $p_{ii}(t)$'s as functions of t is needed in order to choose t' suitably. In this paper we will assume that the series in (4.7) does converge.

Case 2. Some of the parameters q_{ij} are known to be zero. In this case we take as our estimate of Q a modification of $\hat{Q}$. First $\hat{Q}$ is computed by (4.7). Then if $q_{\alpha\beta}$ is known to be 0, $\hat{q}_{\alpha\beta}$ is replaced by 0 and $\hat{q}_{\alpha\alpha}$ altered so that (2.5) is satisfied. The same is done for all $q_{\alpha\beta}$'s known to be 0. Denote the altered matrix by $\hat{\hat{Q}}$. Usually the elements of $\hat{Q}$ and $\hat{\hat{Q}}$ will differ only slightly. However, if the size of $\hat{\hat{Q}}$ is large and there are enough zeros in it, in the right places, then the subsequent computations involving powers of $\hat{\hat{Q}}$ will be considerably lightened.

A trial calculation is sufficient to show that $\hat{\hat{Q}}$ is not, in general the maximum likelihood estimate of Q, subject to the given restrictions. The estimate $\hat{\hat{Q}}$ has been chosen because it is the simplest reasonable estimate in sight. It can be shown that as $t' \to 0$ the relative asymptotic efficiency of the estimate $\hat{\hat{q}}_{ij}$ and the maximum likelihood estimate of q_{ij} converges to 1, for each element $\hat{\hat{q}}_{ij}$ of $\hat{\hat{Q}}$. Therefore, at least for t' small, the possible loss of efficiency in using $\hat{\hat{Q}}$ instead of the maximum likelihood estimate will be small.

On the other hand, $\hat{Q}$ and $\hat{\hat{Q}}$ are both consistent estimates of Q for case 2, so that the large sample distributions of the same rational functions (*e.g.*, $\hat{p}_{ij}(T)$ and $p_{ij}(T)$ defined below) of $\hat{Q}$ and $\hat{\hat{Q}}$ are the same. This follows from the theorems of sections 20.6 and 33.3 of Cramer's ('46) book. We are justified therefore in neglecting the distinction between $\hat{Q}$ and $\hat{\hat{Q}}$ when treating of the large sample distribution properties of such rational functions. We will, in fact, use the notation $\hat{Q}$ for either case 1 or case 2 estimates.

Starting with the estimate $\hat{Q}$ of Q, the estimate of $P(t)$ for any t, $0 \leq t \leq T$, is defined to be

$$\hat{P}(t) = I + \hat{Q}t + \hat{Q}^2 \frac{t^2}{2!} + Q^3 \frac{t^3}{3!} + \cdots \tag{4.8}$$

obtained by substituting $\hat{Q}$ into (2.3). Similarly substituting $\hat{Q}$ into (3.1) we obtain the estimate

$$\hat{e}(T_0) = IT_0 + \hat{Q} \frac{T_0^2}{2!} + Q^2 \frac{T_0^3}{3!} + \cdots \tag{4.9}$$

of $e(T_0)$. If some of the transition rates in $\hat{Q}$ are suppressed giving an estimated net rate matrix of the form $\hat{Q}_{\alpha\beta|\cdots|\mu\nu}$, then the corresponding

net transition and net residence matrices are defined similarly. The suppression of rates in $\hat{Q}$ to obtain $\hat{Q}_{\alpha\beta|\cdots|\mu\nu}$ should not be confused with the suppression of rates in $\hat{Q}$ to obtain $\hat{\hat{Q}}$, the case 2 estimate. The latter is an estimate of Q, while $Q_{\alpha\beta|\cdots|\mu\nu}$ is not.

It will be observed, in forming the matrix products $(\hat{P}^*(t'))^n$ and $\hat{Q}^n$ for $n = 1, 2, \cdots$ that the sums of the elements of the rows of these powers equals 0. This provides a convenient check in computing.

5. THE LARGE SAMPLE DISTRIBUTION OF THE ESTIMATES.

The joint asymptotic distribution of the $\hat{q}_{ij}$'s will be found first. As explained in the previous section, it is sufficient to consider only the case 1 estimates. The only restrictions on these estimates are that $-\hat{q}_{ii} = \sum_{j \neq i} \hat{q}_{ij}$ for $i = 1, \cdots, k$. Therefore, to avoid a singular distribution we consider the joint distribution only of the $\hat{q}_{ij}$'s for $i \neq j$. The distribution of the $\hat{q}_{ij}$'s can be found, if necessary, from the distribution of the $\hat{q}_{ij}$'s, $i \neq j$. The $\hat{q}_{ij}$'s for $i \neq j$ satisfy the regularity conditions of section 33.3 of Cramer's ('46) book; hence we may conclude that these estimates

1. converge in probability to the corresponding elements of Q,

2. are jointly asymptotically normally distributed with the asymptotic covariance matrix equal to the inverse of the matrix

$$\left\{ -\mathcal{E}\left(\frac{\partial^2 \log L}{\partial q_{\alpha\beta} \partial q_{\gamma\delta}}\right)\right\} \qquad \alpha, \beta, \gamma, \delta = 1, \cdots, k;\ \alpha \neq \beta, \gamma \neq \delta$$

where $\mathcal{E}$ denotes "expected value of."

From (4.5)

$$(5.1) \quad -\mathcal{E}\left(\frac{\partial^2 \log L}{\partial q_{\alpha\beta}{}^2}\right) = -\sum_{i=1}^{k} \sum_{j=1}^{k} \mathcal{E}(n_{ij}) \left[-\frac{1}{p_{ij}(t')^2} \left(\frac{\partial p_{ij}(t')}{\partial q_{\alpha\beta}}\right)^2 + \frac{1}{p_{ij}(t')} \left(\frac{\partial^2 p_{ij}(t')}{\partial q_{\alpha\beta}{}^2}\right)\right].$$

To calculate $\mathcal{E}(n_{ij})$, first consider the path of a single patient, starting from time 0 in the initial state and classified subsequently at times $t', 2t', \cdots, st' = T$ the length of the study period. Let x_a be the number of transitions from state S_i at time $(a-1)t'$ to state S_j at time at'. Then the expected number of transitions $S_i \to S_j$ for the

whole path is $\mathcal{E}(x_1+x_2+\cdots+x_s)=\mathcal{E}(x_1)+\cdots+\mathcal{E}(x_s)$. Now

$$\mathcal{E}(x_a)=p_{bi}((a-1)t')p_{ij}(t') \qquad a=1,\cdots,s$$

if S_b is the initial state. Therefore

$$\mathcal{E}(x_1+\cdots+x_s)=\sum_{a=1}^{s}p_{bi}((a-1)t')p_{ij}(t').$$

If there are n_b patients with the initial state S_b for $b=1,\cdots,k$ and n_{ij} is the total number of transitions $S_i\to S_j$ for all the $n=\sum n_b$ patients, then

$$\mathcal{E}(n_{ij})=\sum_{b=1}^{k}n_b\sum_{a=1}^{s}p_{bi}((a-1)t')p_{ij}(t').$$

Substituting in (5.1),

$$\begin{aligned}-\mathcal{E}\left(\frac{\partial^2\log L}{\partial q_{\alpha\beta}{}^2}\right)&=-\sum_{i,j}\left\{p_{ij}(t')\left[\sum_b n_b\sum_{a=0}^{s-1}p_{bi}(at')\right]\right.\\&\qquad\left.\cdot\left[\frac{-1}{p_{ij}{}^2(t')}\left(\frac{\partial p_{ij}(t')}{\partial q_{\alpha\beta}}\right)^2+\frac{1}{p_{ij}(t')}\left(\frac{\partial^2 p_{ij}(t')}{\partial q_{\alpha\beta}{}^2}\right)\right]\right\}\\&=\sum_i\left[\sum_b n_b\sum_{a=0}^{s-1}p_{bi}(at')\right]\left[\sum_j\frac{1}{p_{ij}(t')}\left(\frac{\partial p_{ij}(t')}{\partial q_{\alpha\beta}}\right)^2\right]\\&=\sum_b n_b\left\{\sum_i\left[\sum_{a=0}^{s-1}p_{bi}(at')\right]\left[\sum_j\frac{1}{p_{ij}(t')}\left(\frac{\partial p_{ij}(t')}{\partial q_{\alpha\beta}}\right)^2\right]\right\}.\end{aligned}\tag{5.2}$$

If t' is sufficiently small while T remains fixed, this last expression can be approximated by one of simpler form, for we have

$$\lim_{\substack{t'\to 0\\ st'=T}}t'\sum_{a=0}^{s-1}p_{bi}(at')=e_{bi}(T)$$

$$\lim_{\substack{t'\to 0\\ st'=T}}\sum_j\frac{1}{t'p_{ij}(t')}\left(\frac{\partial p_{ij}(t')}{\partial q_{\alpha\beta}}\right)^2=\begin{cases}\dfrac{1}{q_{\alpha\beta}} & \text{if } i=\alpha\\ 0 & i\neq\alpha\end{cases}$$

using l'Hospital's rule for the second limit. Therefore

$$\lim_{\substack{t'\to 0\\ st'=T}}-\mathcal{E}\left(\frac{\partial^2\log L}{\partial q_{\alpha\beta}{}^2}\right)=\frac{1}{q_{\alpha\beta}}\sum_b n_b e_{ba}(T).\tag{5.3}$$

The expressions (5.2) and (5.3) have been calculated for two systems of transition rates and for several values of t' in order to give

some idea of the effect of t' on the accuracy of the approximation (5.3). The first system is one used as an example by Fix and Neyman. This system has four states with the rate matrix

$$Q = \begin{bmatrix} -3.7 & 1.0 & 2.7 & 0 \\ 0 & 0 & 0 & 0 \\ .3 & 0 & -.4 & .2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The interpretation of the states is of no consequence for this illustration. For this system it is assumed that all patients have the initial state S_1.

The second system was chosen arbitrarily and consists of three states with the rate matrix

$$Q = \begin{bmatrix} -.65 & .45 & .20 \\ 1.20 & -1.97 & .77 \\ .33 & .68 & -1.01 \end{bmatrix}.$$

It is assumed that the initial states are divided equally among the three states.

TABLE 1

Asymptotic Variances of Estimates of Transition Rates for Two Systems of True Values of Rates

TRUE RATES		PROPORTION t' OF TOTAL PERIOD OF OBSERVATION T BETWEEN SUCCESSIVE CLASSIFICATIONS			
		$t' = \frac{1}{6}T$	$t' = \frac{1}{10}T$	$t' = \frac{1}{20}T$	$t' \to 0$
System* I	$q_{12} = 1.0$	$3.59/n$	$3.57/n$	$3.56/n$	$3.56/n$
	$q_{13} = 2.7$	$10.30/n$	$9.92/n$	$9.74/n$	$9.60/n$
	$q_{21} = .2$	$.60/n$	$.52/n$	$.47/n$	$.42/n$
	$q_{24} = .2$	$.42/n$	$.42/n$	$.42/n$	$.42/n$
System** II	$q_{12} = .45$		$.42/n$	$.38/n$	$.35/n$
	$q_{13} = .20$		$.19/n$	$.17/n$	$.16/n$
	$q_{21} = 1.20$		$1.72/n$	$1.63/n$	$1.54/n$
	$q_{23} = .77$		$1.13/n$	$1.06/n$	$.99/n$
	$q_{31} = .33$		$.42/n$	$.38/n$	$.35/n$
	$q_{32} = .68$		$.83/n$	$.77/n$	$.72/n$

* For system I, n is the number of patients with initial state S_1, the only initial state.

** For system II, equal numbers n of patients have the initial states S_1, S_2, and S_3.

We find, similarly, for the asymptotic covariances

$$(5.4)\qquad -\mathcal{E}\left(\frac{\partial^2 \log L}{\partial q_{\alpha\beta}\partial q_{\gamma\delta}}\right) = \sum_b n_b \left\{ \sum_i \left[\sum_{a=0}^{s-1} p_{bi}(at') \right] \cdot \left[\sum_j \frac{1}{p_{ij}(t')} \left(\frac{\partial p_{ij}(t')}{\partial q_{\alpha\beta}}\right)\left(\frac{\partial p_{ij}(t')}{\partial q_{\gamma\delta}}\right)\right]\right\}$$
$$= 0 \text{ as } t' \to 0,\ s \to \infty, \text{ and } st' = T.$$

We conclude that for n large and t' small the $\hat{q}_{\alpha\beta}$'s for $\alpha \neq \beta$ are approximately normally independently distributed with the mean of $\hat{q}_{\alpha\beta}$ equal to $q_{\alpha\beta}$ and the variance equal to

$$(5.5)\qquad \frac{q_{\alpha\beta}}{\sum_b n_b e_{ba}(T)}$$

in this normal distribution.

Estimates of (5.2), (5.4) and (5.5) are obtained by substituting sample estimates for the corresponding parameters. For example, in (5.2), $\hat{p}_{bi}(at')$, $\hat{p}_{ij}(t')$ and $\frac{\partial \hat{p}_{ij}(t')}{\partial \hat{q}_{\alpha\beta}}$ are substituted for $p_{bi}(at')$, $p_{ij}(t')$, and $\frac{\partial p_{ij}(t')}{\partial q_{\alpha\beta}}$ respectively.

Turning now to $\hat{p}_{ij}(t)$ and $\hat{e}_{ij}(t)$ for $0 \leq t \leq T$, it can be shown that these estimates are asymptotically normally distributed. The asymptotic mean of $\hat{p}_{ij}(t)$ is $p_{ij}(t)$ and the asymptotic variance is

$$(5.6)\qquad \frac{1}{n} \sum_{\alpha\neq\beta} \sum_{\gamma\neq\delta} \left(\frac{\partial p_{ij}(t)}{\partial q_{\alpha\beta}}\right)\left(\frac{\partial p_{ij}(t)}{\partial q_{\gamma\delta}}\right) c'_{\alpha\beta\gamma\delta}(n),$$

where the derivatives are evaluated at the true values of $q_{\alpha\beta}$, $q_{\gamma\delta}$, etc., and $\frac{1}{n}\, c'_{\alpha\beta\gamma\delta}(n)$ is the asymptotic covariance of $\hat{q}_{\alpha\beta}$ and $\hat{q}_{\gamma\delta}$. For t' (the interval between successive observations) small this is approximately

$$(5.7)\qquad \sum_{\alpha\neq\beta} \left(\frac{\partial p_{ij}(t)}{\partial q_{\alpha\beta}}\right)^2 \frac{q_{\alpha\beta}}{\sum_b n_b e_{ba}(T)}.$$

The asymptotic mean of $\hat{e}_{ij}(t)$ is $e_{ij}(t)$, while the exact and approximate asymptotic variances, corresponding to (5.6) and (5.7), are

$$(5.8)\qquad \frac{1}{n} \sum_{\alpha\neq\beta} \sum_{\gamma\neq\delta} \left(\frac{\partial e_{ij}(t)}{\partial q_{\alpha\beta}}\right)\left(\frac{\partial e_{ij}(t)}{\partial q_{\gamma\delta}}\right) c'_{\alpha\beta\gamma\delta}(n)$$

and

$$\sum_{\alpha\neq\beta}\left(\frac{\partial e_{ij}(t)}{\partial q_{\alpha\beta}}\right)^2 \frac{q_{\alpha\beta}}{\sum_b n_b e_{b\alpha}(T)}. \tag{5.9}$$

The proof is deferred to appendix 3. Estimates of (5.6), (5.7), (5.8), and (5.9) are found by substituting $\hat{q}_{\alpha\beta}$ for $q_{\alpha\beta}$ throughout.

In computing the derivatives in the preceding formulas, the matrices

$$\left\{\frac{\partial p_{ij}(t)}{\partial q_{\alpha\beta}}\right\}, \left\{\frac{\partial e_{ij}(t)}{\partial q_{\alpha\beta}}\right\}$$

can be computed as follows: For any matrix $A=\{a_{ij}(x)\}$ define the derivative of A with respect to x as the matrix of the derivatives of its elements, *i. e.*,

$$\frac{\partial A}{\partial x}=\left\{\frac{\partial a_{ij}(x)}{\partial x}\right\}.$$

We have then,

$$\frac{\partial P(t)}{\partial q_{\alpha\beta}}=\frac{\partial Q}{\partial q_{\alpha\beta}}t+\left(\frac{\partial Q}{\partial q_{\alpha\beta}}Q+Q\frac{\partial Q}{\partial q_{\alpha\beta}}\right)\frac{t^2}{2!} + \left(\frac{\partial Q^2}{\partial q_{\alpha\beta}}Q+Q^2\frac{\partial Q}{\partial q_{\alpha\beta}}\right)\frac{t^3}{3!}+\cdots. \tag{5.10}$$

Letting

$$\begin{aligned} A_1&=\frac{\partial Qt}{\partial q_{\alpha\beta}}\\ A_2&=\tfrac{1}{2}A_1Qt+\tfrac{1}{2}QtA_1\\ A_3&=\tfrac{1}{3}A_2Qt+\tfrac{1}{3}\left(\frac{Q^2t^2}{2!}\right)A_1\\ A_4&=\tfrac{1}{4}A_3Qt+\tfrac{1}{4}\left(\frac{Q^3t^3}{3!}\right)A_1\\ &\cdots\cdots\cdots \end{aligned}$$

(5.10) becomes

$$\frac{\partial P(t)}{\partial q_{\alpha\beta}}=A_1+A_2+A_3+\cdots. \tag{5.11}$$

We find similarly

$$\frac{\partial e(t)}{\partial q_{\alpha\beta}}=A_1\frac{t}{2}+A_2\frac{t}{3}+A_3\frac{t}{4}+\cdots. \tag{5.12}$$

6. TESTS OF HYPOTHESES

In this section we are concerned with two processes, having the same sets of states but with possibly different rate (and hence different transition and residence) matrices. We shall distinguish between the two processes, in our notation, by means of the superscripts "(1)", and "(2)".

Some questions that might arise in connection with a follow-up study are:

1. Are the rate matrices of two periods the same or not? Such a question might arise in deciding whether to combine two adjacent periods into a single one of length equal to the sum of the two original periods. Or it might arise in deciding whether to combine two sets of data, possibly collected at different places, representing parallel studies.

2. If $p^{(1)}(T)$ and $p^{(2)}(T)$ are transition probabilities for the period $[0, T]$, corresponding to two different treatments (and hence two different processes), are they the same?

3. The same question as the previous one with respect to $e^{(1)}(T)$ and $e^{(2)}(T)$.

Considering question one first, an approximate large sample test of the null hypothesis that the rates $q_{\alpha\beta}^{(1)}$ and $q_{\alpha\beta}^{(2)}$ with $\alpha \neq \beta$ are equal, against the alternative hypothesis that $q_{\alpha\beta}^{(1)} \neq q_{\alpha\beta}^{(2)}$, is based on the statistic

$$U_{\alpha\beta} = \frac{|\hat{q}_{\alpha\beta}^{(1)} - \hat{q}_{\alpha\beta}^{(2)}|}{\sqrt{\sigma^2(\hat{q}_{\alpha\beta}^{(1)}) + \sigma^2(\hat{q}_{\alpha\beta}^{(2)})}} \tag{6.1}$$

where $\sigma^2(\hat{q}_{\alpha\beta}^{(1)})$ and $\sigma^2(\hat{q}_{\alpha\beta}^{(2)})$ are (estimates of) the asymptotic variances of $\hat{q}_{\alpha\beta}^{(1)}$ and $\hat{q}_{\alpha\beta}^{(2)}$ respectively. These are given by (5.2), or (5.3) if t' is small. From section 5 it follows that $U_{\alpha\beta}$ is distributed approximately as the absolute value of a unit normal variate. The hypothesis of the equality of $q_{\alpha\beta}^{(1)}$ and $q_{\alpha\beta}^{(2)}$ is rejected for a large value of $U_{\alpha\beta}$ (depending on the significance level). It will be noted that $T^{(1)}$ need not equal $T^{(2)}$.

For t' small we can also test the equality of several rates jointly. Let the null hypothesis be that

$$q_{\alpha_1\beta_1}^{(1)} = q_{\alpha_1\beta_1}^{(2)}, q_{\alpha_2\beta_2}^{(1)} = q_{\alpha_2\beta_2}^{(2)}, \cdots, q_{\alpha_p\beta_p}^{(1)} = q_{\alpha_p\beta_p}^{(2)},$$

where $\alpha_i \neq \beta_i$ for $i = 1, \cdots, p$ and the (α_i, β_i)'s are distinct, and the alternative hypothesis be that at least one of the equalities fails to hold.

Since the $\hat{q}_{\alpha\beta}$'s are asymptotically independent as $n \to \infty$ and $t' \to 0$, it follows that for large n and small t'

$$\sum_{i=1}^{p} U_{\alpha_i\beta_i}{}^2 = \chi_p{}^2 \tag{6.2}$$

is distributed approximately as χ^2 with p degrees of freedom. In particular, we can test the hypothesis that $Q^{(1)} = Q^{(2)}$ against the alternative that $Q^{(1)} \neq Q^{(2)}$ by summing (6.2) over all (α, β)'s for which $\alpha \neq \beta$. The hypothesis of equality is rejected for a large value of $\chi_p{}^2$ (depending on the significance level).

The significance levels of the tests based on (6.1) and (6.2) should not be taken too literally, as the tests are based on several assumptions: the normal distributions assumed for the $\hat{q}_{\alpha\beta}$'s are asymptotic, not exact; the variances in the denominator of (6.1) are estimates, not exact expressions; and unless $t' = 0$, the $\hat{q}_{\alpha\beta}$'s are only approximately independent, asymptotically. Nothing much more specific can be said about the effect on the significance level without inquiring more deeply into the distributions of the statistics (6.1) and (6.2), or conducting some sampling experiments.

With respect to the elements of the transition and residence matrices for $T^{(1)} = T^{(2)} = T$, statistics similar to (6.1) can be used for large sample tests of hypotheses of the form $p_{\alpha\beta}{}^{(1)}(T) = p_{\alpha\beta}{}^{(2)}(T)$ or $e_{\alpha\beta}{}^{(1)}(T) = e_{\alpha\beta}{}^{(2)}(T)$ for $\alpha \neq \beta$, against either one or two sided alternatives of inequality. For example, to test the hypothesis that $p_{\alpha\beta}{}^{(1)}(T) = p_{\alpha\beta}{}^{(2)}(T)$ against the alternative that $p_{\alpha\beta}{}^{(1)}(T) > p_{\alpha\beta}{}^{(2)}(T)$ we may use the statistic

$$V_{\alpha\beta} = \frac{\hat{p}_{\alpha\beta}{}^{(1)}(T) - \hat{p}_{\alpha\beta}{}^{(2)}(T)}{\sqrt{\sigma^2(\hat{p}_{\alpha\beta}{}^{(1)}(T)) + \sigma^2(\hat{p}_{\alpha\beta}{}^{(2)}(T))}}, \tag{6.3}$$

which is approximately normally distributed with zero mean and unit variance, for large n. The hypothesis of equality is rejected for large positive values of $V_{\alpha\beta}$. Here again $\sigma^2(\hat{p}_{\alpha\beta}{}^{(1)}(T))$ and $\sigma^2(\hat{p}_{\alpha\beta}{}^{(2)}(T))$ are (estimates of) the asymptotic variances of $\hat{p}_{\alpha\beta}{}^{(1)}(T)$ and $\hat{p}_{\alpha\beta}{}^{(2)}(T)$ respectively, and these variances are given by expressions of the form (5.6), or (5.7) if t' is small.

7. TESTS OF GOODNESS OF FIT OF THE MODEL

We have so far dealt with tests of hypotheses assuming that the model fits the disease process adequately. There are certain internal

checks (*i.e.*, checks based on the data alone) of the goodness of fit of the model that can be made. These checks are of conditions 2 and 3 of section 2.

Condition 2 is: the probability that a patient in state S_i at any time during the study, is subsequently in state S_j at a later time in the study, depends on the two states in question but not on the previous history of the patient. As an alternative to condition 2, we may take the hypothesis: if a patient at times t_{-1} and t_0 $(t_{-1} < t_0)$ is in states S_h and S_i, respectively, then the probability of his being in state S_j at time t_1 $(t_0 < t_1; t_{-1}, t_0, t_1$ any times in the period of observation) depends on S_h as well as on S_i and S_j, for at least one (triple S_h, S_i, S_j) of states.

A test of condition 2 against this alternative is the following. Assuming that the interval between successive classifications of n patients is t', let n_{hij} be the number of transitions $S_h \to S_i \to S_j$. Therefore $\sum_h n_{hij} = n_{ij}$. If condition 2 is satisfied, then for each pair (i, j),

$$\hat{p}_{hij}(t') = \frac{n_{hij}}{\sum_{j=1}^{k} n_{hij}} \qquad h = 1, \cdots, k, \tag{7.1}$$

have the same expected values; if the alternative hypothesis is true then for at least one pair (i, j), the expected values of $\hat{p}_{hij}(t')$ for $h = 1, \cdots, k$ will not be the same. Accordingly, we may use a χ^2 test of homogeneity of proportions. It will be observed that if, and only if, q_{ii} is zero, then the proportions (7.1) are either identically one or zero for all samples. Letting E be the set of indices i for which q_{ii} is not *a priori* zero, we have

$$\chi_u^2 = \sum_{h \in E} \sum_{i \in E} \sum_{j=1}^{k} \frac{(n_{hij} - \mathcal{E}n_{hij})^2}{\mathcal{E}n_{hij}} \tag{7.2}$$

$$\mathcal{E}n_{hij} = \frac{(\sum_{j=1}^{k} n_{hij})(\sum_{h \in E} n_{hij})}{\sum_{h \in E} \sum_{j=1}^{k} n_{hij}}.$$

If E has m elements, then n, the degrees of freedom, equals $m[(k-1)m - (k-1)]$. A large value of χ_u^2 renders condition 2 suspect.

The distribution of χ_u^2 is asymptotically χ^2 as $n \to \infty$. For finite n the approximation is poor if any of the n_{hij}'s or $(\sum_j n_{hij})$'s are small. If any of these quantities are small, it will be better to partition the n_{ij}'s into fewer than m parts. Suppose we partition the set E into

$b < m$ subsets, $E_1, E_2, \cdots, E_b$. Define $n_{E_h ij}$ as the number of transitions $S_\alpha \to S_i \to S_j$, where $\alpha \in E_h$. The formulae of (7.2) are then replaced by

$$\chi_v^2 = \sum_{h=1}^{b} \sum_{i \in E} \sum_{j=1}^{k} \frac{(n_{E_h ij} - \mathcal{E} n_{E_h ij})^2}{\mathcal{E} n_{E_h ij}} \tag{7.3}$$

$$\mathcal{E} n_{E_h ij} = \frac{(\sum_{j=1}^{k} n_{E_h ij})(\sum_{h=1}^{b} n_{E_h ij})}{\sum_{h=1}^{b} \sum_{j=1}^{k} n_{E_h ij}}$$

while $v = m[(k-1)b - (k-1)]$.

However, it may happen that no matter how the n_{ij}'s are partitioned there will be a sum $\sum_{h=1} n_{E_h ij}$ or an expectation $\mathcal{E} n_{E_h ij}$ that is too small. In this case, an alternative test can sometimes be used. The indices $1, \cdots, k$ are partitioned into d subsets $(2 \leq d \leq k), F_1, F_2, \cdots, F_d$ and $n_{F_h ij}$ defined to be the number of transitions $S_\alpha \to S_i \to S_j$ where $\alpha \in F_h$. The matrix $\{n_{ij}\}$ is thus partitioned into d matrices $\{n_{F_h ij}\}$, $(h = 1, \cdots, d)$, from which are calculated matrices Q_h, $(h = 1, \cdots, d)$. The Q_h's are then tested for equality pairwise by the test of the previous section. This check involves considerable computation for $d > 2$, however.

If condition 3 is violated, one way in which it is likely to be violated is that some of the elements of Q undergo an increasing or decreasing trend with time. The ordinary life table regarded as a two state process with the states "alive" and "dead" is an example of a process which violates condition 3 in this manner, if we take as the total period of observation the first five years. Here the transition probability from the state "alive" to the state "dead" for one year decreases considerably from the first to the last year. The condition is better met for five years in the middle of the table. As a test of condition 3 against the alternative of a trend, the total period of the study can be divided in half and $n_{ij}^{(1)}$ and $n_{ij}^{(2)}$ computed for the first and second halves respectively for $i \in E$ and $j = 1, \cdots, k$, where E is defined above. Again, we use a χ^2 test of homogeneity with

$$\chi_w^2 = \sum_{h=1}^{2} \sum_{i \in E} \sum_{j=1}^{k} \frac{(n_{ij}^{(h)} - \mathcal{E} n_{ij}^{(h)})^2}{\mathcal{E} n_{ij}^{(h)}} \tag{7.6}$$

$$\mathcal{E} n_{ij}^{(h)} = \frac{(\sum_{j=1}^{k} n_{ij}^{(h)})(\sum_{h=1}^{2} n_{ij}^{(h)})}{\sum_{j=1}^{k} \sum_{h=1}^{2} n_{ij}^{(h)}}.$$

Here $w = m[(k-1)2-(k-1)]$. Alternatively, we can compute $\hat{Q}^{(1)}$ and $\hat{Q}^{(2)}$ for the first and second halves respectively and test these matrices for equality.

What do we do if these tests show that condition 2 or 3 is apparently violated?

If condition 3 is violated, the total period can be divided into sub-periods small enough for the rate matrices of each sub-period to satisfy condition 3 approximately. The method of section 8 is then used to join the sub-periods together.

If condition 2 is violated, it may be possible to substitute a higher order Markov process, *i. e.*, one in which the transition probabilities are determined by two or more preceding states. The generalization of the method of this paper to such a higher order model is straightforward, but will not be done here.

8. TREATMENT OF TWO CONTIGUOUS STUDY PERIODS

Assume that there are k states for each period and that the first period is from $t=0$ to $t=T_1$ and the second from $t=T_1$ to $t=T_2$. For the first period, let Q_1 be the rate matrix,

$$P_1(t) = \{p_{ij}^{(1)}(t)\} \qquad 0 \leq t \leq T_1$$

the corresponding transition matrix, and

$$(8.1) \qquad e_1(t') = \{e_{ij}^{(1)}(t')\} = \int_0^{t'} P_1(t)\,dt \qquad 0 \leq t' \leq T_1$$

the residence matrix. For the second period, let the corresponding matrices be Q_2,

$$P_2(t-T_1) = \{p_{ij}^{(2)}(t+T_1)\} \qquad T_1 \leq t \leq T_2$$

and

$$(8.2) \qquad e_2(t'-T) = \{e_{ij}^{(2)}(t'-T_1)\} = \int_{T_1}^{t'} P_2(t-T_1)\,dt \qquad T_1 \leq t' \leq T_2.$$

It follows that if at time 0 an individual is in state S_i, the probability $p_{ij}(t)$ of being in state S_j at time t is

$$p_{ij}(t) = \begin{cases} p_{ij}^{(1)}(t) & 0 \leq t \leq T_1 \\ \sum_{a=1}^{k} p_{ia}^{(1)}(T_1)\, p_{aj}^{(2)}(t-T_1) & T_1 \leq t \leq T_2. \end{cases}$$

In matrix notation, this is

$$P(t) = \begin{cases} P_1(T) & 0 \leq t \leq T_1 \\ P_1(T_1)P_2(t-T_1) & T_1 \leq t \leq T_2. \end{cases} \tag{8.3}$$

In terms of Q_1 and Q_2,

$$P(t) = \begin{cases} I + Q_1 t + Q_1^2 \frac{t^2}{2!} + \cdots & 0 \leq t \leq T_1 \\ I + [Q_1 T_1 + Q_2(t-T_1)] + [Q_1 T_1 + Q_2(t-T_1)]^2 \frac{1}{2!} & \\ \quad + [Q_1 T_1 + Q_2(t-T_1)]^3 \frac{1}{3!} + \cdots & T_1 \leq t \leq T_2. \end{cases} \tag{8.4}$$

From (8.1), (8.2) and (8.3), we have, for $T_1 \leq t' \leq T_2$,

$$\begin{aligned} e(t') &= \int_0^{t'} P(t)\,dt = \int_0^{T_1} P(t)\,dt + \int_{T_1}^{t'} P(t)\,dt \\ &= \int_0^{T_1} P_1(t)\,dt + P_1(t) \int_{T_1}^{t'} P_2(t-T_1)\,dt \\ &= e_1(T_1) + P_1(T_1)e_2(t'-T_1). \end{aligned} \tag{8.5}$$

As in the case of a single period of study, these formulae hold also for the estimates $\hat{Q}_1, \hat{Q}_2, \hat{P}_1(t), \cdots$ or net estimates

$$\hat{Q}_{1\cdot\alpha\beta|\gamma\delta|\ldots}, \hat{Q}_{2\cdot\alpha\beta|\gamma\delta|\ldots}, \hat{P}_{1\cdot\alpha\beta|\gamma\delta|\ldots}(t), \text{etc.}$$

The results of section 4 concerning the large sample distribution of the $\hat{p}_{ij}(t)$'s and $\hat{e}_{ij}(t)$'s extend without any difficulty to the present case, and we may state that for any pair of indices $i, j = 1, \cdots, k$, $\hat{p}_{ij}(t)$, $T_1 \leq t \leq T_2$ is asymptotically normally distributed with asymptotic mean $p_{ij}(t)$ and asymptotic variance

$$\begin{aligned} &\frac{1}{n^{(1)}} \sum_{\alpha \neq \beta} \sum_{\gamma \neq \delta} \left(\frac{\partial p_{ij}(t)}{\partial q_{\alpha\beta}^{(1)}}\right)\left(\frac{\partial p_{ij}(t)}{\partial q_{\gamma\delta}^{(1)}}\right) c'_{\alpha\beta\gamma\delta^{(1)}}(n^{(1)}) \\ &\quad + \frac{1}{n^{(2)}} \sum_{\alpha \neq \beta} \sum_{\gamma \neq \delta} \left(\frac{\partial p_{ij}(t)}{\partial q_{\alpha\beta}^{(2)}}\right)\left(\frac{\partial p_{ij}(t)}{\partial q_{\gamma\delta}^{(2)}}\right) c'_{\alpha\beta\gamma\delta^{(2)}}(n^{(2)}). \end{aligned} \tag{8.6}$$

For t' small we may use approximations of the form (5.7). The formula (8.6) is the generalization of (5.6).

The matrix

$$\frac{\partial P(t)}{\partial q_{\alpha\beta}^{(h)}} \qquad T_1 \leq t \leq T_2,\ h = 1, 2$$

can be found from (5.11) using the relations

$$\frac{\partial P(t)}{\partial q_{\alpha\beta^{(1)}}} = \frac{\partial P_1(T_1)}{\partial q_{\alpha\beta^{(1)}}} P_2(t - T_1)$$

$$\frac{\partial P(t)}{\partial q_{\gamma\delta^{(2)}}} = P_1(T_1) \frac{\partial P_2(t - T_1)}{\partial q_{\gamma\delta^{(2)}}}$$

derived from (8.3).

Similarly $\hat{e}_{ij}(t)$ for $T_1 \leq t \leq T_2$ and $i, j = 1, \cdots, k$ is asymptotically normally distributed with asymptotic mean $e_{ij}(t)$ and with the asymptotic variance given by an expression of the same form as (8.6) but with $e_{ij}(t)$ replacing $p_{ij}(t)$. The same remark as before applies if t' is small. Furthermore

$$\frac{\partial e(t)}{\partial q_{\alpha\beta^{(h)}}} \qquad T_1 \leq t \leq T_2,\ h = 1, 2$$

is computed using (5.11) and (5.12) and the relations

$$\frac{\partial e(t)}{\partial q_{\alpha\beta^{(1)}}} = \frac{\partial e_1(t)}{\partial q_{\alpha\beta^{(1)}}} + \frac{\partial P_1(T_1)}{\partial q_{\alpha\beta^{(1)}}} e_2(t - T_1)$$

$$\frac{\partial e(t)}{\partial q_{\gamma\delta^{(2)}}} = P_1(T_1) \frac{\partial e_2(t - T_1)}{\partial q_{\gamma\delta^{(2)}}}$$

derived from (8.5).

An estimate of (8.6) is obtained in the usual way by substituting estimates of the p's and q's for these parameters throughout.

9. EXAMPLE AND COMPARISON OF FIX AND NEYMAN'S AND OUR ESTIMATES

We consider in this section a single period only. It was not possible to obtain any actual data with which to illustrate the estimates of this paper. Instead, samples were generated artificially based on a system of parameters (already given in section 5) used by Fix and Neyman ('51). To generate the samples, the Fisher and Yates ('49) tables of random digits, and the Rand Corporation series of random digits appearing in recent issues of the *Journal of the American Statistical Association* were used.

The samples so obtained enable us to compare the estimates of Fix and Neyman with ours to give an indication of the relative performances of the estimates. As a basis for comparison, the sample variances of

the estimates of the risks, for the two estimates in question, have been used. As the estimates of this paper are based on somewhat more information than Fix and Neyman's estimates, it is to be expected that the sample variances of our estimates will tend to be less. The question of greater interest is how much less; unfortunately the present experiment is too small to answer this question.

The example that we use has the rate matrix

$$\begin{bmatrix} -3.7 & 1.0 & 2.7 & 0 \\ 0 & 0 & 0 & 0 \\ .2 & 0 & -.4 & .2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Briefly, the interpretation is:

S_1: active cancer

S_2: death from cancer or operative death

S_3: cured or apparently well

S_4: lost from observation or death from other causes.

The *a priori* assumptions are that $q_{14} = q_{22} = q_{32} = q_{44} = 0$. Letting the total period of observation equal 1, and observing the "patients" at intervals of 1/10, the true transition matrix for the interval 1/10 is, by (2.3),

$$P(1/10) = \begin{bmatrix} .693 & .084 & .221 & .002 \\ 0 & 1 & 0 & 0 \\ .016 & .001 & .963 & .020 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

On the basis of $P(1/10)$, a total of 3000 paths were generated, half with the initial state S_1 and half with the initial state S_3. For example, the first four of these paths are:

1	1	1	3	3	3	3	3	3	3
1	1	3	3	3	4	4	4	4	4
1	1	3	3	3	3	3	3	3	3
1	1	1	1	1	1	1	1	1	1.

The 3000 paths were divided up into 20 groups. The first ten groups each contained 50 paths with initial state S_1 and 50 with initial state S_3; the second 10 groups each contained 100 paths with initial state S_1 and 100 with initial state S_3. The paths were subdivided into these groups

according to their order of generation. Following the notation of section 4, table 2 gives the numbers of transitions $S_i \rightarrow S_j$ for each sample. It will be observed that the remaining n_{ij}'s are *a priori* 0 or 1.

The *a priori* restrictions on Q are such that q_{12}, q_{13}, q_{31}, and q_{34} determine Q, and therefore it is sufficient to give the estimates of these parameters only. Our estimates of these parameters are based on (4.7),

TABLE 2

Numbers of Transitions in Twenty Samples

SAMPLE	n_{11}	n_{12}	n_{13}	n_{14}	n_{31}	n_{33}	n_{33}	n_{34}
1	124	16	37	0	10	0	638	13
2	165	13	40	0	8	1	636	11
3	142	15	40	0	8	1	611	15
4	137	22	29	0	5	1	597	11
5	140	13	36	0	6	0	620	16
6	126	12	40	1	6	1	666	15
7	114	12	37	0	6	0	642	16
8	153	13	43	0	12	0	609	15
9	110	16	36	1	4	0	619	15
10	140	16	37	0	10	2	609	17
11	238	44	66	2	19	1	1170	33
12	236	38	79	0	23	0	1246	24
13	231	24	89	1	23	2	1339	20
14	235	31	78	1	16	0	1273	29
15	271	35	81	0	26	1	1241	24
16	288	25	83	2	24	1	1300	22
17	232	20	85	0	13	1	1303	35
18	239	34	74	2	19	1	1286	22
19	293	28	80	1	20	0	1286	24
20	255	26	81	1	15	2	1264	30

while Fix and Neyman's estimates are based on the formulae of page 230 of their paper. It will be noted that their state S_i is our state S_{i+1} for $i = 0, 1, 2, 3$. The estimates for the 20 samples are given in table 3.

Fix and Neyman's estimates for samples 6 and 12 are infinite and have therefore been omitted.

In table 4 are given the sample variances of the transition rates based on the two groups of samples, 1 to 10, and 11 to 20. It will be seen from table 4 that the estimates of this paper, with one exception, have the smaller variances.

In practice, one would also be interested in the matrices $P(1)$ and $e(1)$, and therefore in the variances of the estimates of these two

matrices. It can be shown, however, that the variances of the $\hat{p}_{ij}(1)$'s and $\hat{e}_{ij}(1)$'s are approximately equal to linear combinations, with positive coefficients, of the variances of the $\hat{q}_{ij}$'s, so that for an indication of the performance of the estimates in general, it is sufficient to examine the sample variances of the $\hat{q}_{ij}$'s.

TABLE 3

Estimates of Transition Rates: (A) *Ours,* (B) *Fix and Neyman's*

	ESTIMATES OF							
	$q_{12} = 1.0$		$q_{13} = 2.7$		$q_{21} = .2$		$q_{24} = .2$	
SAMPLE*	A	B	A	B	A	B	A	B
1	1.08	1.06	2.54	2.21	.18	.17	.20	.13
2	.68	.79	2.14	2.04	.14	.00	.17	.22
3	.89	1.20	2.43	2.71	.15	.05	.24	.38
4	1.36	1.34	1.83	1.88	.10	.08	.18	.14
5	.80	.67	2.25	1.64	.11	.06	.25	.16
6	.79	—	2.70	—	.11	—	.22	—
7	.88	.72	2.75	2.10	.11	.15	.26	.13
8	.73	.96	2.45	2.95	.23	.47	.24	.04
9	1.19	1.28	2.72	2.63	.08	.00	.24	.36
10	.97	.94	2.30	2.28	.19	.35	.27	.01
11	1.52	1.37	2.33	2.13	.19	.21	.28	.29
12	1.31	—	2.78	—	.22	—	.19	—
13	.84	1.02	3.19	3.58	.21	.12	.15	.20
14	1.08	1.12	2.78	2.79	.15	.21	.22	.16
15	1.08	1.40	2.55	3.21	.24	.25	.19	.13
16	.74	.84	2.49	3.07	.21	.44	.17	—.02
17	.71	.60	3.09	2.62	.12	.04	.26	.33
18	1.17	1.05	2.59	2.49	.18	.18	.17	.11
19	.81	.84	2.37	2.38	.18	.13	.18	.23
20	.85	1.00	2.70	2.91	.14	.10	.23	.19

* Samples 1-10 consist of 100 paths, half with initial state S_1 and half with initial state S_3. Samples 11-20 have twice as many paths in each category.

Besides $P(1)$ and $e(1)$ the other matrices of interest in this example are $P_{34}(1)$ and $e_{34}(1)$, which are the transition and residence matrices with the "inessential" risk q_{34} suppressed. The estimates of $P(1)$, $e(1)$, $P_{34}(1)$, and $e_{34}(1)$ have been computed from (2.3) and (3.1), only for samples 1 and 10, and only for the estimates of this paper. These estimates are given in table 5. Together with estimates of their variances obtained from (5.7) and (5.9), they are examples of solutions of the estimation problem in our model.

TABLE 4

Sample Variances about True Values, of Estimates of Transition Rates, Based on Two Sample Sizes of (A) *Our Estimates and* (B) *Fix and Neyman's Estimates*

SAMPLE VARIANCE OF	ESTIMATE	SAMPLE SIZE = TOTAL NUMBER OF PATHS IN EACH SAMPLE		NUMBER OF SAMPLES
		100	200	
$\hat{q}_{12}$	A	.046	.064	10
	B	.052	.058	9
$\hat{q}_{13}$	A	.160	.073	10
	B	.340	.192	9
$\hat{q}_{21}$	A	.006	.002	10
	B	.026	.012	9
$\hat{q}_{24}$	A	.002	.002	10
	B	.015	.010	9

TABLE 5

True Values and Our Estimates of $P(1)$, $e(1)$, $P_{24}(1)$, *and* $e_{24}(1)$

ESTIMATE OF	SAMPLE 1	SAMPLE 10	WITH RATE q_{24} SUPPRESSED ESTIMATE OF *	SAMPLE 1	SAMPLE 10
$p_{11}(1) = .054$	.053	.043	$p_{11.24}(1) = .036$	.057	.047
$p_{12}(1) = .281$	.307	.406	$p_{12.24}(1) = .324$	.308	.408
$p_{13}(1) = .570$	.549	.444	$p_{13.24}(1) = .641$	.636	.545
$p_{14}(1) = .095$	.091	.106			
$p_{21}(1) = .042$	.040	.036	$p_{21.24}(1) = .047$	.046	.045
$p_{22}(1) = .035$	.035	.048	$p_{22.24}(1) = .086$	.038	.054
$p_{23}(1) = .751$	.752	.687	$p_{23.24}(1) = .866$	.916	.902
$p_{24}(1) = .172$	.173	.228			
$e_{11}(1) = .281$	.285	.268	$e_{11.24}(1) = .162$	.286	.269
$e_{12}(1) = .205$	.223	.301	$e_{12.24}(1) = .255$	.224	.302
$e_{13}(1) = .474$	.453	.385	$e_{13.24}(1) = .583$	.490	.429
$e_{14}(1) = .040$	.038	.046			
$e_{21}(1) = .035$	.033	.032	$e_{21.24}(1) = .043$	.035	.035
$e_{22}(1) = .015$	.015	.021	$e_{12.24}(1) = .039$	.016	.022
$e_{23}(1) = .860$	.862	.827	$e_{23.24}(1) = .918$	.949	.942
$e_{24}(1) = .090$	.090	.121			

* $p_{ij.24}(1)$ is the ij-th element of $P_{24}(1)$; similarly for the residence matrix.

In table 5 we see that sample 1 happened to yield estimates that are quite close to the true values, whereas sample 10 with twice as many observations has not done as well. In practice not all of the estimates in such a table as 5 may have a useful interpretation. In our case, $e_{12}(1)$, for example, would have little practical significance.

It may be of interest to know that to compute $P(1/10)$, Q, $P(1)$, $e(1)$, $P_{34}(1)$, and $e_{34}(1)$ for a single sample took about 6 hours. A 4 state model with no *a priori* restrictions would have taken about twice as long. It took about 15 terms of the series (2.3) and (3.1) to secure convergence accurate to 5 decimal places. To compute Q took fewer terms.

APPENDIX 1

Theorem: If

$$Qt = P^*(t) - \frac{(P^*(t))^2}{2} + \frac{(P^*(t))^3}{3} - \cdots \tag{A1.1}$$

converges (elementwise) then (A1.1) is equivalent to

$$P^*(t) = Qt + Q^2\frac{t^2}{2!} + Q^3\frac{t^3}{3!} + \cdots. \tag{A1.2}$$

Proof: The coefficients of (A1.2) are the same as the coefficients of the power series representation of $e^x - 1 = y$:

$$y = x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots, \tag{A1.3}$$

while the coefficients of (A1.1) are the same as the coefficients of the power series representation of $x = \log(1+y)$, for $|y| < 1$:

$$x = y - \frac{y^2}{2} + \frac{y^3}{3} - \cdots. \tag{A1.4}$$

First assume (A1.2) is true. Substituting (A1.2) into (A1.1) we obtain on the right side of (A1.1) the series

$$Qt + a_2(Qt)^2 + a_3(Qt)^3 + \cdots \tag{A1.5}$$

where $a_2, a_3, \cdots$ are certain constants. The series (A1.5) converges by hypothesis. Substituting (A1.3) into (A1.4) we obtain on the right side of (A1.4) the series

$$x + a_2x^2 + a_3x^3 + \cdots \tag{A1.6}$$

which has the identical coefficients as (A1.5). This follows from the fact that $P^*(t)$, Qt, x and y enter only as algebraic elements following the same multiplication and addition laws. However, since (A1.6) converges and equals x for an infinite number of values of x, all the coefficients $a_2, a_3, \cdots$ are zero. Therefore (A1.5) equals Qt and (A1.1) is true. The argument for the converse is the same given the fact that the exponential series converges for all values of the argument.

APPENDIX 2

A sufficient condition for the convergence (elementwise) of

$$\sum_{n=1}^{\infty} P^{*(n)}(t)\frac{(-1)^{n+1}}{n} \tag{A2.1}$$

is that

$$p_{ii}(t) > \tfrac{1}{2} \quad \text{for} \quad i = 1, \cdots, k. \tag{A2.2}$$

Proof: Denote the ij'th element of $P^{*(n)}(t)$ by $p^*{}_{ij}{}^{(n)}(t)$. (A2.2) is equivalent to $\sum_{a \neq i} p_{ia}(t) < \frac{1}{2}$ for $i = 1, \cdots, k$; hence $|p^*{}_{ij}{}^{(1)}(t)| < \frac{1}{2}$ for $i, j = 1, \cdots, k$. Now assume that

$$|p^*{}_{ij}{}^{(n)}(t)| < \tfrac{1}{2} \quad \text{for} \quad i, j = 1, \cdots, k \tag{A2.3}$$

for some value of n. By definition,

$$\begin{aligned} p^*{}_{ij}{}^{(n+1)}(t) &= \sum_{a=1}^{k} p^*{}_{ia}(t)p^*{}_{aj}{}^{(n)}(t) \\ &= \sum_{a} p_{ia}(t)p^*{}_{aj}{}^{(n)}(t) - p^*{}_{ij}{}^{(n)}(t) \\ &= \sum_{a \neq i} p_{ia}(t)p^*{}_{aj}{}^{(n)}(t) + (p_{ii}(t) - 1)p^*{}_{ij}{}^{(n)}(t). \end{aligned}$$

By (A2.2) and (A2.3) the absolute values of each of the last two summands is less than $\frac{1}{4}$, so that $|p^*{}_{ij}{}^{(n+1)}(t)| < \frac{1}{2}$. Hence, by induction, (A2.3) holds for all n.

Now, by (A2.2), there exists a $\delta > 1$ such that $|\delta p^*{}_{ii}(t)| < \frac{1}{2}$ for $i = 1, \cdots, k$. Substituting $\delta p^*{}_{ij}(t)$ for $p^*{}_{ij}(t)$, for all i, j, it follows from what has just been proved that

$$|\delta^n p^*{}_{ij}{}^{(n)}(t)| < \tfrac{1}{2} \quad \text{or} \quad |p^*{}_{ij}{}^{(n)}(t)| < \frac{1}{2\delta^n}$$

for $i, j = 1, \cdots, k$ and all n. Hence,

$$\sum_{n=1}^{\infty} | p^{*}{}_{ij}{}^{(n)}(t) \frac{(-1)^{n+1}}{n} | < \tfrac{1}{2} \sum_{n=1}^{\infty} \frac{1}{n\delta^n} \quad \text{for} \quad i, j = 1, \cdots, k.$$

The infinite sum on the right converges, and since absolute convergence implies convergence, it follows that (A2.1) converges.

APPENDIX 3

We give here the derivation of the asymptotic distributions of $\hat{p}_{ij}(t)$ and $\hat{e}_{ij}(t)$.

Observe first that (5.4) has the form

$$n \sum_{b} \frac{n}{n_b} h_b(\alpha, \beta, \gamma, \delta) = n c_{\alpha\beta\gamma\delta}(n)$$

where $n = n_1 + \cdots + n_k$ and $c_{\alpha\beta\gamma\delta}(n)$ depends on n but is bounded as $n \to \infty$. Therefore, the asymptotic covariance of $\hat{q}_{\alpha\beta}$ and $\hat{q}_{\gamma\delta}$ for $\alpha \neq \beta$ and $\gamma \neq \delta$ has the form

$$\text{(A3.1)} \qquad \frac{1}{n} c'_{\alpha\beta\gamma\delta}(n)$$

where $c'_{\alpha\beta\gamma\delta}(n)$ depends on n but is bounded. In particular $\hat{q}_{\alpha\beta}$ has the asymptotic variance $\frac{1}{n} c'_{\alpha\beta\alpha\beta}(n) = \frac{1}{n} c'_{\alpha\beta}{}^{2}(n)$.

Now expanding $\hat{p}_{ij}(t)$ in a Taylor expansion about the $q_{\alpha\beta}$'s, we obtain

$$\text{(A3.2)} \qquad \sqrt{n}(\hat{p}_{ij}(t) - p_{ij}(t)) = \sqrt{n} \sum_{\alpha \neq \beta} \left(\frac{\partial p_{ij}(t)}{\partial q_{\alpha\beta}} \right) (\hat{q}_{\alpha\beta} - q_{\alpha\beta}) + \sqrt{n} R.$$

The remainder R is the quadratic form

$$\text{(A3.4)} \qquad \sum_{\alpha \neq \beta} \sum_{\gamma \neq \delta} \left(\frac{\partial^2 p'_{ij}(t)}{\partial q'_{\alpha\beta} \partial q'_{\gamma\delta}} \right) (\hat{q}_{\alpha\beta} - q_{\alpha\beta}) (\hat{q}_{\gamma\delta} - q_{\gamma\delta}),$$

where the derivatives are evaluated at points $q'_{\alpha\beta}, q'_{\gamma\delta}, \cdots$ intermediate to $\hat{q}_{\alpha\beta}$ and $q_{\alpha\beta}$, $\hat{q}_{\gamma\delta}$ and $q_{\gamma\delta}, \cdots$ respectively.

We show first that $\sqrt{n}R$ converges in probability to zero. Since $\hat{q}_{\alpha\beta}$ is asymptotically normally distributed, we have

$$\text{(A3.5)} \qquad \Pr\{ | \hat{q}_{\alpha\beta} - q_{\alpha\beta} | \geq \frac{1}{\sqrt{n}} c'_{\alpha\beta}(n) u \} \to 2(1 - \Phi(u))$$

where $\Phi(u)$ is the unit normal distribution function. Let ϵ and δ be given such that $1 > \epsilon > 0$, and $1 > \delta > 0$. First choose u so large that

$$2(1-\Phi(u)) < \frac{\delta}{k(k-1)} \tag{A3.5}$$

(where k is the size of the Q matrix). Then choose n' such that for $n > n'$ all the inequalities

$$\frac{1}{\sqrt[4]{n}} c'_{\alpha\beta}(n)u < \epsilon \qquad \alpha \neq \beta \tag{A3.7}$$

are satisfied simultaneously. Finally, for each n, define Z_n as the set of sample points for which all the inequalities

$$|\hat{q}_{\alpha\beta} - q_{\alpha\beta}| < \frac{1}{\sqrt{n}} c'_{\alpha\beta}(n)u \qquad \alpha \neq \beta \tag{A3.8}$$

are satisfied simultaneously, and define Z^*_n as the complementary set.

It follows from (A3.5) and A3.6) that for all n greater than some n'', say,

$$\Pr(Z^*_n) < k(k-1)\frac{\delta}{k(k-1)} = \delta$$

and hence $\Pr(Z_n) \geq 1 - \delta$ for $n > n''$. But on Z_n it follows from (A3.7) and (A3.8) that for $n > n'$,

$$\sqrt{n}\,|\hat{q}_{\alpha\beta} - q_{\alpha\beta}|\,|\hat{q}_{\gamma\delta} - q_{\gamma\delta}| < \frac{1}{\sqrt{n}} c'_{\alpha\beta}(n)c'_{\gamma\delta}(n)u^2$$

$$< \frac{1}{\sqrt{n'}} c'_{\alpha\beta}(n)c'_{\gamma\delta}(n)u^2 < \epsilon^2.$$

Since ϵ and δ are arbitrarily close to zero, it follows that the term on the left converges in probability to zero, and hence $\sqrt{n}R$, a linear combination of such terms, also converges in probability to zero.

Now consider $\sqrt{n}\,|\hat{q}_{\alpha\beta} - q_{\alpha\beta}|$. We have

$$\Pr\{\sqrt{n}\,|\hat{q}_{\alpha\beta} - q_{\alpha\beta}| \geq c'_{\alpha\beta}(n)u\} \to 2(1-\Phi(u)).$$

As $u \to 0$, $2(1-\Phi(u)) \to 1$, so that $\sqrt{n}\,|\hat{q}_{\alpha\beta} - q_{\alpha\beta}|$ does not converge in probability to zero.

It follows, cancelling $\sqrt{n}$, that $(\hat{p}_{ij}(t) - p_{ij}(t))$ has the same asymptotic distribution as

$$\sum_{\alpha\neq\beta} \left(\frac{\partial p_{ij}(t)}{\partial q_{\alpha\beta}}\right)(\hat{q}_{\alpha\beta} - q_{\alpha\beta}).$$

This is a linear combination of asymptotically normal variables, hence is itself asymptotically normal, with the asymptotic mean 0 and asymptotic variance

$$\frac{1}{n}\sum_{\alpha\neq\beta}\sum_{\gamma\neq\delta}\left(\frac{\partial p_{ij}(t)}{\partial q_{\alpha\beta}}\right)\left(\frac{\partial p_{ij}(t)}{\partial q_{\gamma\delta}}\right)c'_{\alpha\beta\gamma\delta}(n).$$

For t', (the interval between successive observations) sufficiently small, this is approximately

$$\sum_{\alpha\neq\beta}\left(\frac{\partial p_{ij}(t)}{\partial q_{\alpha\beta}}\right)^2\frac{q_{\alpha\beta}}{\sum_b n_b e_{b\alpha}(T)}.$$

The asymptotic distribution of the $\hat{e}_{ij}(t)$'s is derived similarly. $\hat{e}_{ij}(t)$ is asymptotically normal with asymptotic mean $e_{ij}(t)$ and asypmtotic variance

$$\frac{1}{n}\sum_{\alpha\neq\beta}\sum_{\gamma\neq\delta}\left(\frac{\partial e_{ij}(t)}{\partial q_{\alpha\beta}}\right)\left(\frac{\partial e_{ij}(t)}{\partial q_{\gamma\delta}}\right)c'_{\alpha\beta\gamma\delta}(n).$$

For t' sufficiently small this is approximately

$$\sum_{\alpha\neq\beta}\left(\frac{\partial e_{ij}(t)}{\partial q_{\alpha\beta}}\right)^2\frac{q_{\alpha\beta}}{\sum_b n_b e_{b\alpha}(T)}.$$

SUMMARY

The general, time stationary, continuous, Markov process with a finite number of states is introduced as a model for a follow-up study. This model may prove useful in situations in which several states of the disease can be distinguished and are of clinical importance, as opposed to just two states, life, and death. A special case of this process was used by Fix and Neyman ('51) for a 4 state model.

The estimate of the parameters of the model, called transition rates, is based on observations of patients at equal intervals of time t', for a total period of time $st' = T$. Letting $P(t')$ be the matrix of transition probabilities for $t = t'$, and $\hat{P}(t')$ the maximum likelihood estimate of $P(t')$, the matrix Q of transition rates is estimated by

$$\hat{Q} = \log \hat{P}(t') = (\hat{P}(t') - I) - (\hat{P}(t') - I)^2/2 + (\hat{P}(t') - I)^3/3 - \cdots,$$

provided the series converges. A simple condition for convergence is that the diagonal elements of $\hat{P}(t') - I$ be less than $\frac{1}{2}$.

The asymptotic distributions of the elements of $\hat{Q}$, $\hat{P}(t)$, and

$\hat{e}(t) = \int_0^t \hat{P}(u)du$, the residence matrix, are all normal. The effect of letting $t' \to 0$, upon the asymptotic variance of the elements of $\hat{Q}$ is examined in a special case. When t' is small the formulas simplify considerably.

Some tests of hypotheses concerning the equality of transition rates, transition probabilities, and residencies (elements of $e(t)$), and some goodness of fit tests of the model are given. The model is extended to two (or more) contiguous periods, each with separate transition rate matrices. Finally the results of a sampling experiment using a 4 state model are given.

ACKNOWLEDGEMENT

I wish to thank Dr. William E. Reynolds of the Harvard Medical School for keeping me in touch with the clinician's viewpoint of the subject matter of this paper.

LITERATURE CITED

CRAMER, H. 1946 Mathematical Methods of Statistics. Princeton University Press, Princeton.

DOOB, J. L. 1953 Stochastic Processes. John Wiley and Son, Inc., New York.

FELLER, W. 1950 An Introduction to Probability Theory and Its Applications. Vol. 1. John Wiley and Son, Inc., New York, 307-362.

FISHER, R. A. AND F. YATES 1948 Statistical Tables for Biological Agricultural and Medical Research, 3rd ed. Oliver and Boyd, Edinburgh, 104-109.

FIX, E., AND J. NEYMAN 1951 A simple stochastic model of recovery, relapse, death, and loss of patients. Human Biology, *23*: 205-241.

WEDDERBURN, J. H. M. 1934 Lectures on Matrices. American Mathematics Society Colloquium Publications, 122-123.

NOTE ON ESTIMATION IN A MARKOV PROCESS WITH CONSTANT TRANSITION RATES *

BY PAUL MEIER

Department of Biostatistics, The Johns Hopkins University

REMARKS

IN the foregoing paper Mr. Zahl presents as a model for follow-up studies the general Markov process with a finite number of states, say $S_1, \cdots, S_k$, and constant transition rates, q_{ij} = transition rate from S_i to S_j. As he remarks, special cases have been examined before, notably by Fix and Neyman ('51). Indeed, perhaps the simplest case of this process is just the familiar exponential survivorship (or constant death rate) model so often used both in medical and industrial applications.

In both the Fix-Neyman paper and in the foregoing excellent study, observation is assumed to be incomplete—the status of each subject is known only at certain times during the follow-up interval. Although such restrictions are clearly appropriate to many applications, there are others in which transitions from one state to another are sufficiently dramatic (*e.g.*, living to dead) that observation may be regarded as essentially complete over the follow-up interval. Since, also, the maximum likelihood estimates of the basic parameters, q_{ij}, are exceedingly simple when there is complete information, it seems appropriate to introduce them along with the general stochastic model.

* Paper no. 304 from the Department of Biostatistics.

THE ESTIMATES

Briefly, q_{ij} is the constant rate of transition from state S_i to state S_j and the maximum likelihood estimate of q_{ij} is simply

$$\hat{q}_{ij} = \frac{m_{ij}}{T_i},$$

where

m_{ij} = total transitions observed from S_i to S_j, summed over all subjects,

T_i = total time under observation in state S_i, summed over all subjects.

This is a well known result for the special case of exponential survivorship mentioned earlier (Littell, '52, and Epstein and Sobel, '53) in which we have one transient state (surviving) and one absorbing state (dead or failed). The extension of the derivation to the general case is straightforward and will not be discussed here.

AN EXAMPLE

Consider a model for chronic disease follow-up with three states, S_1, S_2, S_3, representing the states "well," "sick," and "dead." (S_1 and S_2 are "transient" states and S_3 is an "absorbing" state.) Suppose the study involves only two individuals starting in S_2, with the following histories. The first becomes well at 1.4 years, relapses at 2.3 years, regains health at 3.2 years and remains well for the rest of the 5-year follow-up interval. The second becomes well at 0.8 years, relapses at 1.6 years, and dies at 3.0 years. The data may be tabulated as shown in table 1.

TABLE 1

CASE 1		CASE 2	
state	period	state	period
S_2	0.0 to 1.4	S_2	0.0 to 0.8
S_1	1.4 to 2.3	S_1	0.8 to 1.6
S_2	2.3 to 3.2	S_2	1.6 to 3.0
S_1	3.2 to 5.0	S_1	3.0 to 5.0

Since S_3 is an absorbing state, $m_{31} \equiv m_{32} \equiv 0$, and T_3 is not needed. The required quantities are

$$T_1 = [(2.3 - 1.4) + (5.0 - 3.2)] + [(1.6 - 0.8)]$$
$$= 2.7 + 0.8 = 3.5$$

$$T_2 = [(1.4 - 0.0) + (3.2 - 2.3)] + [(0.8 - 0.0) + (3.0 - 1.6)]$$
$$= 2.3 + 2.2 = 4.5$$

$$m_{12} = 1 + 1 = 2 \qquad m_{13} = 0 + 0 = 0$$

$$m_{21} = 2 + 1 = 3 \qquad m_{23} = 0 + 1 = 1$$

The estimates of the transition rates are

$$\hat{q}_{12} = 2/3.5 = 0.571 \qquad \hat{q}_{13} = 0/8.1 = 0.000$$

$$\hat{q}_{21} = 3/4.5 = 0.667 \qquad \hat{q}_{23} = 1/4.5 = 0.222.$$

DISTRIBUTION OF ESTIMATES

One's intuition about the distribution of these estimates may be sharpened by examining some methods for generating a random observation from any given process. The method indicated in Doob ('53) is as follows. A subject starting in state S_{i_0} remains there for a period of time randomly chosen from an exponential distribution with mean $1/q_{i_0}$ $(q_{i_0} = \sum_{j \neq i_0} q_{i_0 j})$. He is then transferred to a state S_{i_1}, chosen at random from the states other than S_{i_0}, the chance of going to S_j being $q_{i_0 j}/q_{i_0}$ $(j \neq i_0)$. The subject then remains in state S_{i_1} a time randomly chosen from an exponential distribution with mean $1/q_{i_1}$ and then is transferred to a state different from S_{i_1} and so on. An equivalent method, suggested by the notion of competition among the states, is to choose $k - 1$ times at random from exponential distributions with means $1/q_{i_0 j}$ $(j \neq i_0)$ and make a transfer to that state, S_{i_1}, for which the chosen time was least, making the transition at this minimum time. (The needed random observations from exponential distributions can be generated from random digits, utilizing the fact that if x is uniformly distributed in the interval $(0, 1)$, $-q \ln x$ has an exponential distribution with mean q.)

The exact distribution of the q_{ij} is not at all a simple matter. If they could be regarded as being generated by *fixed and predetermined* total observation times, T_i, the m_{ij} would have independent Poisson distributions and the $\hat{q}_{ij}$ would have mean q_{ij} and variance q_{ij}/T_i, giving a coefficient of variation of $\dfrac{\sqrt{q_{ij}/T_i}}{q_{ij}} = \dfrac{1}{\sqrt{q_{ij} T_i}} = \dfrac{1}{\sqrt{E\{m_{ij}\}}}$. In fact, the

likelihood function does not distinguish between the cases of fixed and random T_i so that the asymptotic formula for the variance of $\hat{q}_{ij}$, derived in the usual way, is just q_{ij}/T_i. Unfortunately the T_i *are* random variables, and the above results for the mean and variance are therefore only approximate. If, however, T_i is large compared to the mean time required for one transfer from S_i to S_j, that is $1/q_{ij}$, the effect of its variability should be slight. Since T_i must be many times $1/q_{ij}$ for $\hat{q}_{ij}$ to have even modest precision, it seems reasonable to expect that the asymptotic Poisson theory will give good service in the cases of most interest, that is, when the data provide reasonably precise estimates of the q_{ij}.

Although the hypothetical study used in the example is far too small to provide useful estimates, it will serve to exhibit the method of calculation of the variance estimates. Letting $\hat{\sigma}_{ij}^2 = \hat{q}_{ij}/T_i$, the estimated variance of $\hat{q}_{ij}$, we have

$$\hat{\sigma}_{12}^2 = \frac{0.571}{3.5} = 0.163 \qquad \hat{\sigma}_{13}^2 = \frac{0.000}{3.5} = 0.000$$

$$\hat{\sigma}_{21}^2 = \frac{0.667}{4.5} = 0.148 \qquad \hat{\sigma}_{23}^2 = \frac{0.222}{4.5} = 0.049.$$

The estimated coefficients of variation are given by $1/\sqrt{m_{ij}}$ and range from 58% for $\hat{q}_{21}$ to infinity for $\hat{q}_{31}$, showing the inadequacy of such a sample for estimating rates.

LITERATURE CITED

DOOB, J. L. 1953 Stochastic Processes. John Wiley and Sons, New York.

EPSTEIN, BENJAMIN AND MILTON SOBEL 1953 Life-testing. Journal of the American Statistical Association, *48*: 486-502.

FIX, EVELYN AND J. NEYMAN 1951 A simple stochastic model of recovery, relapse, death and loss of patients. Human Biology, *23*: 205-241.

LITTELL, A. S. 1952 Estimation of the *T*-year survival rate from follow-up studies over a limited period of time. Human Biology, *24*: 87-116.

LONGEVITY OF RETIRED ARMY AND AIR FORCE OFFICERS *

BY C. A. McMAHAN, JOHN K. FOLGER ** AND THOMAS R. FORD

Air Force Personnel and Training Research Center (OERL)
Maxwell Air Force Base, Alabama

INTRODUCTION

DEMOGRAPHERS, actuaries, and vital statisticians have long used average expectation of life as a measure of mortality and of the extent to which man has gained control over his biological environment. Accompanying the rapid industrialization of our economy and the emergence of a distinct labor force in recent years, life table principles have also been applied to computations of average length of working life. Partially as a result of the demographic, economic, and social changes of the last century and particularly of the last 50 years, the number of persons reaching retirement age has increased tremendously and there is now a considerable span of life for millions of persons between the date of retirement and death. This particular study is focused upon the computation of life expectancy after working life, that is, with measuring length of life in a retired status. Few life tables have been constructed for retired populations.

The life table applied to retired populations presents certain problems, both of construction and interpretation. The several values used in its computation can, however, be profitably utilized to: (a) predict future deaths in a retired population for a reasonable period of time,

* Personal views or opinions expressed or implied in this article are not to be construed as necessarily carrying the official sanction of the Department of the Air Force or of the Air Research and Development Command.

** Formerly a member of the staff of the Human Resources Research Institute, now research associate, Southern Regional Education Board, Atlanta, Georgia.

The authors are indebted to Jerry Walker Combs, Jr., for critically reading the manuscript and for valuable suggestions.

(b) measure the man-years of life remaining for retired individuals—an index of manpower potential, and (c) provide a stationary population for comparative purposes. The latter is perhaps the most limited in utility, and both (a) and (b) above must be used with due regard for the time reference of the life table functions and the structural peculiarities of the actual retired population.

MATERIALS AND METHODS

Selection of the Method Used in This Study

The retired population which is the concern of this study was composed of 10,001 United States Army (and Air Force) officers living in retirement between January 1, 1925 and December 31, 1948. From this group, three were omitted for lack of complete information, leaving 9,998 for analysis. Of the latter, 1,906 were living in retirement at the beginning of the period, and 8,092 retired on or after January 1, 1925. During the 24-year period there were 2,795 deaths, 1,509 of which occurred among officers retiring during the period. Just above 7,000 (7,203) were therefore living in retirement on December 31, 1948.

The primary consideration for the selection of the methodology used was that the population was small and changing rapidly in both size and composition. Because of the small number and limited period of time covered, a cohort life table could not be constructed (Jaffe, '51, p. 3). Such a table would at any rate have had the added disadvantage of the varying ages at which individuals entered the population and the purely historical interest of the mortality experience reflected.[1] On the other hand, the retired population at any one time was too small to allow the construction of a current life table by the usual method of using a mid-point population and a three-year average of deaths.

The method selected may perhaps be classified as a current multiple decrement table, but certain departures in selection of population and use of deaths have been made (Dublin *et al.*, '49, pp. 279-280 and U. S. Dept. of Labor, '50). Instead of being based on a specified calendar year, the rates computed are for a hypothetical "average" year in this period. Each of the retired officers has been counted in the population at each year of age or portion of year during which his life experience

[1] Observed survival ratios have been computed, however, for those who were in a retired status when the study began as well as for those who retired during the period 1925-1929.

could be observed. Thus, most individuals entered into several age groups as a result of aging during the 24-year span. Conceivably an officer could have entered into the base population of 24 different age groups. In actuality, the average number was 7 or 8.[2] Thus the population for any year of age consisted of all officers in retirement who attained that age during the period and all officers who retired at that age. The 50-year age group, for example, consisted of all officers who retired at age 50 or attained age 50 in retirement during the 24-year period. The 51-year age group consisted of all officers who survived their fiftieth year, plus those who retired at age 51. In general, since additional retirements occurred as age increased, and because officers retiring at younger ages grew older during the period, the population used for computing age-specific rates generally increased up to age 64. It was only after that age that attrition through death outweighed the increments to the population.

Obviously, it was not possible to treat the population arrived at as an enumerated population. It was possible, however, to follow two courses. The population could have been computed on the basis of the total number attaining to each x birthday in retirement, making possible the direct computation of the probability of dying; or it was possible to convert the population into years lived in retirement at age x, making possible the computation of the death rate. Because the former would require sacrificing certain data and would involve complications in making the computations by 5-year age groups, and because of certain advantages in age-specific death rates themselves, it was decided to compute the death rate and to deal with the population on a years-lived (years of exposure) basis.

Accordingly, the ${}_nm_x$ (a convenient summary of the conventional notation used can be found in Reed and Merrell, '39) for any given age x was computed from $\frac{{}_nD_x}{{}_nL'_x}$ rather than from the more usual $\frac{{}_nD_x}{{}_nP_x}$.[3] The ${}_nD_x$ included all of the deaths for an age group occurring during the 24-year span, the ${}_nL'_x$ representing the years lived (observed) in retire-

[2] Computed by dividing total man-years lived during the period by the maximum population, 9,998. This gives the average years lived in retirement.

[3] Deaths in the age group x to $x+n = {}_nD_x$; population in the age group x to $x+n = {}_nP_x$. For convenience, the observed person-years lived at each age has been denoted by ${}_nL'_x$. Thus the death rate ${}_nm_x$ really amounts to deaths per 1,000 years of exposure.

ment in the interval x to $x+n$. It is immediately apparent that the formula ${}_nm_x = \frac{{}_nD_x}{{}_nL'_x}$ is roughly equivalent to deriving the ${}_nm_x$ from the $\frac{{}_nd_x}{{}_nL_x}$ values of a life table (${}_nd_x$ = deaths in the life table age group x to $x+n$; ${}_nL_x$ = life table population in the age group x to $x+n$). Theoretically it may be considered a more accurate value for ${}_nm_x$ than that usually employed, but it is perhaps influenced at ages up to 65 by the fact that increments to the population occurred during a year of life, so that both the numerator and denominator are expanding values during a year of life. The theoretical consequences of this fact have not been thoroughly explored. They undoubtedly have some bearing on the convertibility of the ${}_nm_x$ values to ${}_nq_x$ (probability of dying) values by the Reed and Merrell method.

The methods used in actual computations were those adaptable to (1) data available and (2) IBM tabulation.[4]

Available Data

The following data were available for use in the study of length of life of retired officers (Official Registers, '25-'49):

Recorded data. (1) Date of birth of each officer, (2) date of retirement of each officer, (3) date of death of each officer who died, (4) date of beginning of study (January 1, 1925), (5) date of termination of study (December 31, 1948).

Computed data. (1) Age at retirement of each officer, (2) age as of December 31, 1924 for those retiring previous to January 1, 1925, (3) age at death of each officer who died, (4) age as of December 31, 1948 for those officers alive on that date. (All ages were computed in years to two decimal places.) The various possibilities of contributions to the years lived in retirement were listed and are shown in figure 1.

Years Lived in Retirement (Years of Exposure)

The years lived in retirement were computed by the method illustrated in table 1; this method allowed a detailed check of the data at a

[4] The authors are indebted to Stephen J. Zolczynski for furnishing the IBM methodology; they are indebted to Stephen W. Fotis for providing a manual verification on a sample basis of the method used.

COMPUTATION OF AGGREGATE YEARS LIVED IN RETIREMENT, BY AGE

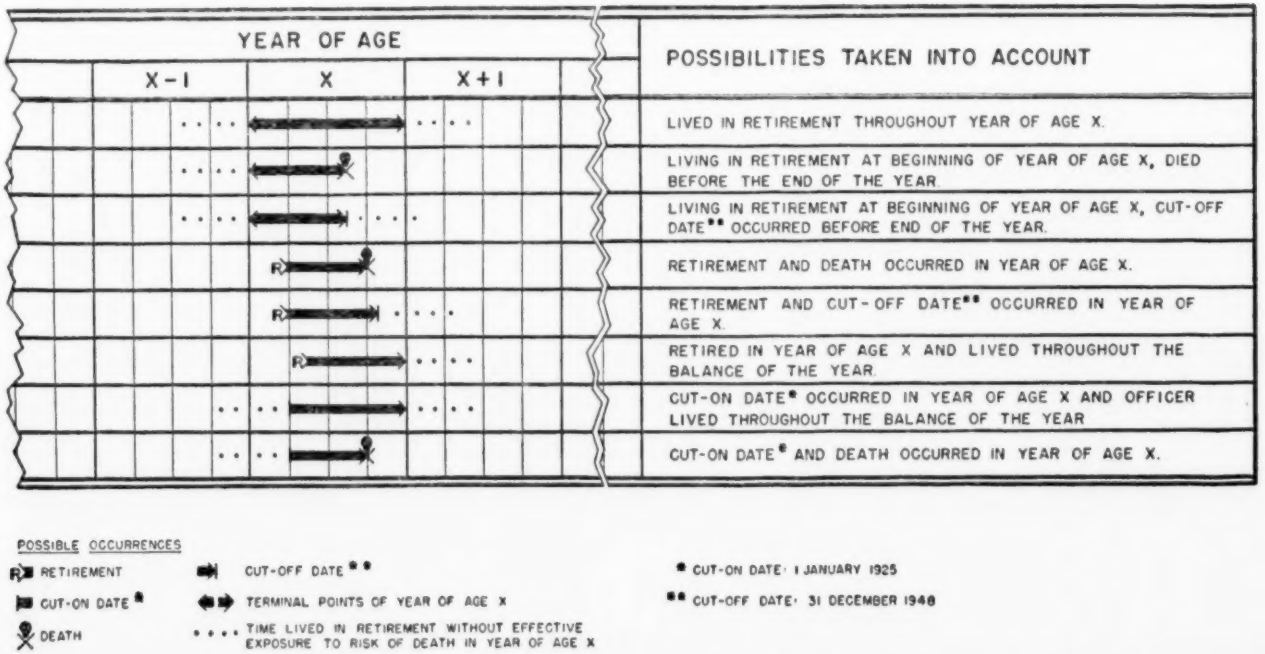
YEAR OF AGE
X-1
X
X+1
POSSIBILITIES TAKEN INTO ACCOUNT
LIVED IN RETIREMENT THROUGHOUT YEAR OF AGE X.
LIVING IN RETIREMENT AT BEGINNING OF YEAR OF AGE X, DIED BEFORE THE END OF THE YEAR.
LIVING IN RETIREMENT AT BEGINNING OF YEAR OF AGE X, CUT-OFF DATE** OCCURRED BEFORE END OF THE YEAR.
RETIREMENT AND DEATH OCCURRED IN YEAR OF AGE X.
RETIREMENT AND CUT-OFF DATE** OCCURRED IN YEAR OF AGE X.
RETIRED IN YEAR OF AGE X AND LIVED THROUGHOUT THE BALANCE OF THE YEAR.
CUT-ON DATE* OCCURRED IN YEAR OF AGE X AND OFFICER LIVED THROUGHOUT THE BALANCE OF THE YEAR.
CUT-ON DATE* AND DEATH OCCURRED IN YEAR OF AGE X.
POSSIBLE OCCURRENCES
R RETIREMENT
CUT-ON DATE*
DEATH
CUT-OFF DATE**
TERMINAL POINTS OF YEAR OF AGE X
TIME LIVED IN RETIREMENT WITHOUT EFFECTIVE EXPOSURE TO RISK OF DEATH IN YEAR OF AGE X
* CUT-ON DATE: 1 JANUARY 1925
** CUT-OFF DATE: 31 DECEMBER 1948

FIG. 1. POSSIBILITIES TAKEN INTO ACCOUNT IN COMPUTING PERSON-YEARS LIVED IN RETIREMENT BY RETIRED OFFICERS OF THE UNITED STATES ARMY AND AIR FORCE, 1925-1948.

glance and yet the machine process was simple.[5] Figure 2 shows the "population" (years of exposure) obtained.

Age Specific Death Rates

Having secured the population it was only necessary to compute ${}_nm_x$ by the formula $\frac{{}_nD_x}{{}_nL'_x}$, as previously given, ${}_nL'_x$ denoting the number of years lived in the interval x to $x+n$. As stated previously, an approximate equivalence between $\frac{{}_nD_x}{{}_nP_x}$, $\frac{{}_nD_x}{{}_nL'_x}$, and $\frac{{}_nd_x}{{}_nL_x}$ has been assumed, although each implies a different type of base population. The ${}_nm_x$ values for the entire retired population, compared with those of white

[5] As the first step in the computation process (IBM), the population was ordered by age at retirement. The next step was to cumulate for each age the number of officers who lived in retirement at that age. Each person who retired was counted as contributing a year at the age of retirement and each succeeding year through age 103. (Age 103 was taken as the final age since no officer in the population lived beyond that age during the period observed. The maximum population of 9,998 was reached at age 94 when the last person had retired.) This gave the total number of retirement years that would have been contributed at each age if (a) a person contributed a full year at the age which he entered the population and (b) no persons left the population. Since these two conditions did not hold, appropriate corrections had to be made. To correct for the fact that few or no officers actually contributed a full year in the year of entry, the actual proportion of the year which they did contribute was computed. The machine technique selected for doing this was to compute and tabulate the proportion of the year which was not contributed. Then by a clerical operation this amount was subtracted from the full year which had been credited at age of entry.

A similar technique was used for subtracting dropouts. Since an officer was credited with a full year for every age after he had entered the population, it was necessary to subtract a full year for every age after he had died and for all ages which would have been attained after December 31, 1948. This was a matter of accumulating the number of dropouts from age 23 onward and subtracting the cumulative total at each age. For the age at which an individual was lost to the population, however, a portion of the year was actually lived in retirement. Since he was already credited with contributing a full year at that age the portion of that year not lived in retirement was subtracted as a correction factor.

In summary, the method was to reckon the maximum retirement years that an officer could contribute and then to subtract all the full years and fractions of years that he was known not to have contributed. Two IBM machine runs only were required, the final number of years lived at each age being obtained by simple subtraction.

TABLE 1

Example of the Computations Used to Obtain Person-Years Lived in Retirement by Retired Officers of the United States Army and Air Force, 1925-1948

	ACCESSIONS				SEPARATIONS				NET
Year of age (1)	Number of persons retiring (IBM) (2)	Cumulative increments —maximum possible person-years of retirement (IBM) (3)	Non-contributive person-years of increments (IBM) (4)	Contributive potential person-years of retirement (3)-(4) (5)	Number of persons dropping out of retired population (IBM) (6)	Cumulative decrements —maximum possible person-years lost through decrements (IBM) (7)	Contributive person-years of decrements (IBM) (8)	Non-contributive years of decrements (7)-(8) (9)	Person-years lived in a retired status (5)-(9) (10)
22	3	3	2.54	.46	—	—	—	—	.46
23	3	6	1.09	4.91	1	1	.48	.52	4.39
24	12	18	6.60	11.40	1	2	.20	1.80	9.60
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
90	1	9993	.02	9992.98	17	9953	7.58	9945.42	47.56
91	1	9994	.81	9993.19	9	9962	3.84	9958.16	35.03
92	3	9997	1.33	9995.67	17	9979	7.79	9971.21	24.46
93	1	9998	.38	9997.62	3	9982	1.32	9980.68	16.94
94	—	9998	—	9998.00	3	9985	1.74	9983.26	14.74
95	—	9998	—	9998.00	3	9988	1.46	9986.54	11.46
96	—	9998	—	9998.00	3	9991	1.44	9989.56	8.44
97	—	9998	—	9998.00	3	9994	2.05	9991.95	6.05
98	—	9998	—	9998.00	1	9995	.19	9994.81	3.19
99	—	9998	—	9998.00	—	9995	—	9995.00	3.00
100	—	9998	—	9998.00	2	9997	1.63	9995.37	2.63
101	—	9998	—	9998.00	—	9997	—	9997.00	1.00
102	—	9998	—	9998.00	1	9998	.36	9997.64	.36
103	—	9998	—	9998.00	—	9998	—	9998.00	—

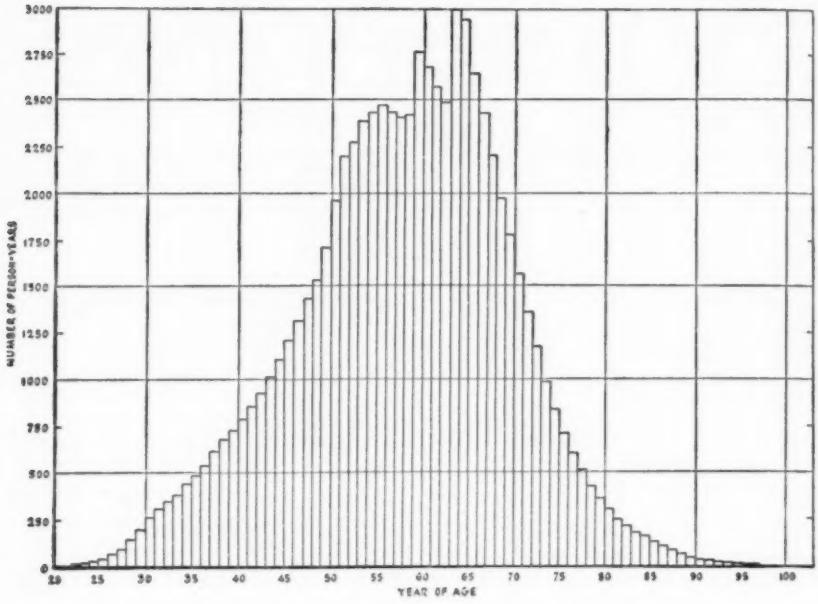

FIG. 2. PERSON-YEARS LIVED IN RETIREMENT BY RETIRED OFFICERS OF THE UNITED STATES ARMY AND AIR FORCE, 1925-1948.

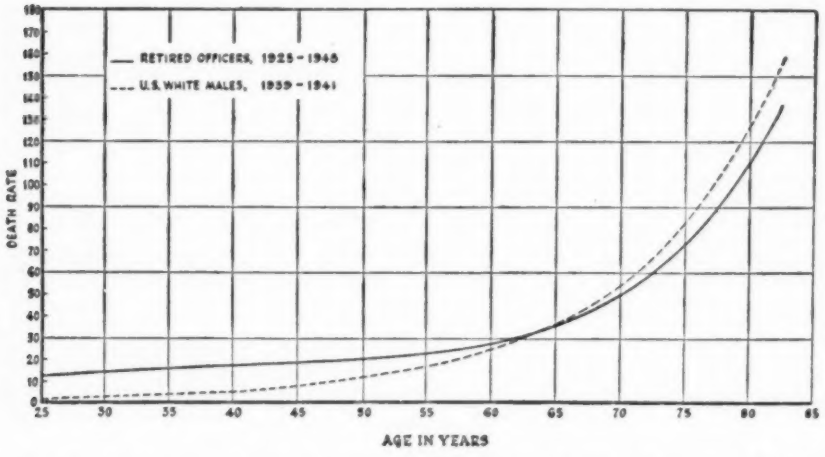

FIG. 3. AGE SPECIFIC DEATH RATES OF WHITE MALES OF THE UNITED STATES, 1939-1941, COMPARED WITH THOSE OF RETIRED OFFICERS OF THE UNITED STATES ARMY AND AIR FORCE, 1925-1948.

males in the United States 1939-1941, are shown in figure 3. The ${}_nm_x$ values computed for specified portions[6] of the retired population are given in figure 4.

Smoothing the data. The smooth curves of mortality applying to retired officers shown in figures 3 and 4 are, of course, not the actual observed rates obtained. Despite the inflation of the population through computing years of exposure, the population at all ages was still extremely small. A total of 10,001 individual officers constituted the initial population, of whom three had to be eliminated because of lack

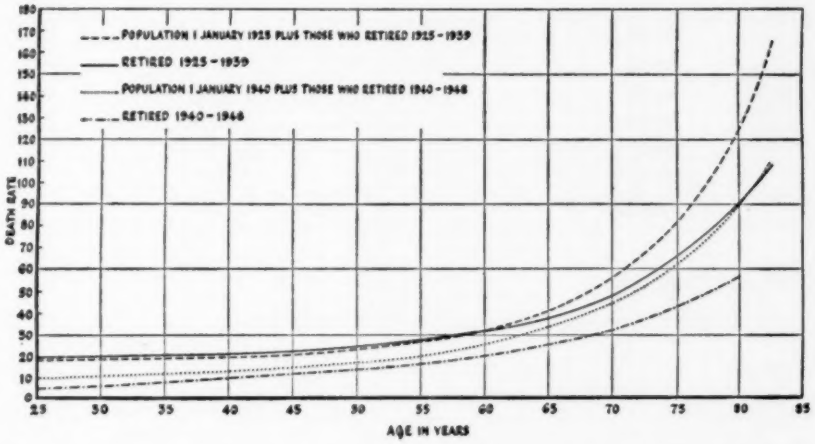
POPULATION 1 JANUARY 1925 PLUS THOSE WHO RETIRED 1925-1939
RETIRED 1925-1939
POPULATION 1 JANUARY 1940 PLUS THOSE WHO RETIRED 1940-1948
RETIRED 1940-1948
DEATH RATE
AGE IN YEARS

FIG. 4. AGE-SPECIFIC DEATH RATES OF SUB-GROUPS OF THE RETIRED OFFICER POPULATION OF THE UNITED STATES ARMY AND AIR FORCE, 1925-1948.

of data. Thus 9,998 officers lived in a retired status averaging less than 8 years each. The total population (person-years lived) at all ages was therefore less than 80,000, actually 76,172.81. Against these there were only 2,795 deaths. The resulting mortality rates were therefore highly irregular. On the usual assumption that these irregularities are not a feature of the "true" mortality rate, they have been graduated mechanically. In this case it appeared that any mathematical formula would be as arbitrary a method of smoothing as drawing the curve with the aid of a spline.

The resultant curves were tested for smoothness by the method of

[6] Computations for these sub-groups were made in the manner previously described by means of additional "cut-on" and "cut-off" dates.

third differences. The results compared favorably with Greville's test of white males for 1939-1941 (Greville, '46, pp. 126-129). In addition, the graduated rates were applied to the actual years of exposure and produced the results shown in table 2. These were considered satisfactory.

TABLE 2

Actual Deaths Compared to Number of Deaths Produced by Smoothed Data for Selected Age Groups of Retired Officers of the United States Army and Air Force, 1925-1948

AGE GROUP	OBSERVED DEATHS	SMOOTHED DATA
30–39	67	71
40–49	230	210
50–59	549	565
60–69	915	924
70–79	685	685
80–89	288	289
Total 30–89	2734	2739

Probability of Dying

Conversion of the computed ${}_nm_x$ values to ${}_nq_x$ by means of the formula $q_x = \frac{2m_x}{2 + m_x}$ was explored but was discarded in favor of the widely-used Reed and Merrell tables ('39). One possible limitation should be recognized here. Since the Reed and Merrell tables give the relationship of ${}_nm_x = \frac{{}_nD_x}{{}_nP_x}$ to ${}_nq_x = \frac{{}_nd_x}{{}_nl_x}$, it has been assumed in this study that the same relationship holds when ${}_nm_x = \frac{{}_nD_x}{{}_nL'_x}$. Some question regarding the equivalence of the two formulas for ${}_nm_x$ has already been raised above, but the Reed and Merrell tables appear more satisfactory for all ranges of ages than the formula $\frac{2m_x}{2 + m_x}$.

RESULTS

The findings of the present study are presented in figures 3 and 4, and in table 3. For the retired population, age-specific rates were higher at ages under 65 than were rates for corresponding ages among the white male population in 1939-1941. Above age 65, however, rates were lower for the retired officer population (figure 3). The higher

TABLE 3

Abridged Life Table for all Retired Officers of the United States Army and Air Force, 1925-1948 (Constructed by Reed and Merrell Method)

AGE INTERVAL	PROPORTION DYING	OF 100,000 RETIRING		STATIONARY POPULATION		AVERAGE REMAINING LIFETIME
Period of life between two exact ages stated in years (1)	Proportion of persons alive at beginning of age interval dying during interval (2)	Number alive at beginning of age interval (3)	Number dying during age interval (4)	In the age interval (5)	In this and all subsequent age intervals (6)	Average number of years of life remaining at beginning of age interval (7)
x to $x+n$	${}_nq_x$	l_x	${}_nd_x$	${}_nL_x$	T_x	$\mathring{e}_x$
25–29	0.06732	100,000	6,732	505,430	3,493,786	34.9
30–34	.07340	93,268	6,846	449,270	2,988,356	32.0
35–39	.08036	86,422	6,945	414,757	2,539,086	29.4
40–44	.08682	79,477	6,900	380,119	2,124,329	26.7
45–49	.09461	72,577	6,867	345,688	1,744,210	24.0
50–54	.10279	65,710	6,754	311,678	1,398,522	21.3
55–59	.11761	58,956	6,934	277,599	1,086,844	18.4
60–64	.14397	52,022	7,490	241,684	809,245	15.6
65–69	.18796	44,532	8,370	202,119	567,561	12.7
70–74	.25806	36,162	9,332	157,791	365,442	10.1
75–79	.36752	26,830	9,861	109,332	207,651	7.7
80–84	.50345	16,969	8,543	62,603	98,319	5.8
85–89	.66606	8,426	5,612	26,777	35,716	4.2
90–94	.77924	2,814	2,193	7,529	8,939	3.2
95–99	.86327	621	536	1,410	1,410	2.3
100–over	—	85	—	—	—	—

mortality for retired officers at younger ages probably reflects (1) the earlier base period for the retired officer rates and (2) the "reasonably expected" lesser longevity of officers retiring at younger ages (Reed and Love, '31). How much of the difference can be attributed to the latter is not known, but there seems to be rather clear evidence that the time reference has influenced the difference. In figure 4 it can be observed that mortality rates are lower for officers retiring recently than for those retiring earlier. The factors effecting an increasing longevity for the general population in recent years thus appear also to be operative among this retired population.

DISCUSSION

When the age-specific death rates are carried forward in the life table to statements of life expectancy, a certain fiction is inevitable. All life tables incorporate such a fiction, since the stationary population is never identical with the actual population and the mortality rates are rarely applicable over the entire lifetime of any actual cohort. The present life table population, however, departs from the actual in rather extreme degree. To what extent this departure results in unreasonable life expectancies at the younger ages, however, is not known. The life expectancy of persons retiring at age 25–29 and earlier has been determined not only by their experience, but by those retiring at older ages. If those retiring at younger ages enjoy materially less longevity than those who retire at older ages, then the life expectancy value for age 25–29 probably misrepresents what could be expected from limiting analysis to this group alone.[7] In projections of total years lived in retirement the values should, of course, be obtained from actual data on persons who retire at such young ages. The life table population is manifestly not appropriate for such purposes. Useful projections may be made, however, by relating the average life expectancy at each age to the number of officers who retire at those ages. As long as the life table values may be considered as approximately applicable to a retired population, estimates of cost or of total years lived in retirement by a retired population may be obtained.

[7] It has been possible to observe 8 individuals retiring at ages 25–29 for a period of 20 years. At the end of that time, 6 were still alive, two had died.

CONCLUSIONS

The technique of life table construction utilized in this study of retired officers might be profitably applied to many groups. For example, in the face of the current teacher shortage, a life table could aid in determining the additional personnel that would be available if the legal age of retirement were delayed five years. It should be realized however that this particular method of life table construction has both advantages and disadvantages. Its chief advantages are as follows: (1) It is an application of a method of life table construction to a relatively small and changing population that can be observed over a period of years when no cohort for a single year is large enough to warrant the construction of a conventional current table or can be observed long enough to make the construction of a cohort life table feasible. (2) It utilizes an uncomplicated machine technique which greatly simplifies the computational procedures. Its limitations, however, must be admitted. The mortality experience is not so recent as that used in strictly current life tables, but is less obsolescent than that used in cohort life tables. The farther the mid-point of the period from the current date, the less predictive value the life table will likely have. Finally, retired populations of the present type have a greater range of ages at which individuals enter the population than do most populations for which life tables have been constructed. Retired officer populations may be more extreme in this regard than other retired populations. Wherever similar populations are encountered, however, average life expectancy values must be interpreted with care.

LITERATURE CITED

DUBLIN, LOUIS I., ALFRED J. LOTKA AND MORTIMER SPIEGELMAN 1949 Lenth of Life. Ronald Press Co., New York.

GREVILLE, THOMAS N. E. 1946 United States Life Tables and Actuarial Tables, 1939-1941. Government Printing Office, Washington.

JAFFE, A. J. 1951 Handbook of Statistical Methods for Demographers. Government Printing Office, Washington.

OFFICIAL ARMY (AND AIR FORCE) REGISTER, ANNUAL VOLUMES. 1925-1949 Government Printing Office, Washington.

REED, LOWELL J., AND ALBERT G. LOVE 1931 Biometric Studies on United States Army Officers. Department of Biostatistics, the Johns Hopkins University, Baltimore.

REED, LOWELL J., AND MARGARET MERRELL 1939 A short method for constructing an abridged life table. Amer. J. Hygiene, *30*: 33-62.

UNITED STATES DEPARTMENT OF LABOR 1950 Tables of Working Life. Government Printing Office, Washington.

THE HEIGHTS AND WEIGHTS OF MEN ACCORDING TO MONTH OF BIRTH

BY A. B. FITT

Auckland University College

AT the time my book on seasonal fluctuations (Fitt, '41) was written, no data had been forthcoming to test the relationship of height and weight of adults to month of birth. Many studies were available of children of different ages. At about the same time two articles appeared offering measurements of American university and college students in a similar kind of analysis. One was by Mills ('41) and the other by Chenoweth and Canning ('41). The trends noted were in many respects similar to those of the author's analysis but not identical. It must be noted that college students represent a highly selected population, not calculated to reveal the growth condition of the population as a whole. In any case they do not represent full adult growth. The height and weight records were taken for 16-, 17-, 18- and 19-year-old freshmen who were still growing (and at different rates).

In order to secure data for a fully adult population it was decided to analyse the heights and weights of a large sample of the men of the New Zealand army draft of World War II. Access was given for this purpose to the Army Base Records, Wellington. To secure as perfect a sample as possible, names were drawn from files classified under each letter of the alphabet and in number proportional to the total number of entries under each letter. The names drawn from each letter-group (which followed a complete alphabetical sequence) were the ones, up to the number required, which came first. It would appear that in this way an almost perfectly random sample of the male population was secured. Height measurements of 21,342 and weight measurements of 21,177 men were thus available.

The analysis of the height measurements when arranged according to the months of birth follows in table 1 which gives the means in inches

along with their probable errors and the number of cases. A more detailed analysis is presented in table 2, a frequency table giving a break-down by months. The original calculations were based on half-inch intervals but for reasons of space inch intervals were used here. In following the months it must be remembered that the data are for the Southern Hemisphere and that the seasons are therefore reversd.

TABLE 1

Heights by Birth-Months

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sep.	Oct.	Nov.	Dec.
Mean (inches)	68.19	68.25	68.09	68.06	68.10	67.94	68.16	68.18	68.24	68.16	68.24	68.17
P.E.*m*	.0409	.0415	.0392	.0389	.0406	.0390	.0392	.0414	.0391	.0406	.0423	.0400
number	1752	1634	1849	1763	1802	1799	1832	1720	1886	1861	1659	1785

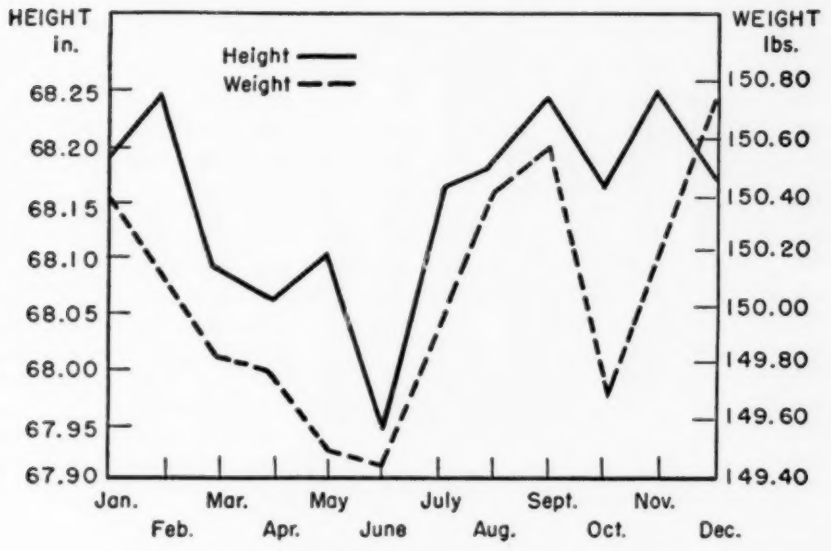

FIG. 1. HEIGHT AND WEIGHT BY BIRTH MONTHS.

Table 1 shows the tallest men were born in February and the shortest in June. Those of September and November birth are just shorter than those born in February. The differences between the June mean and those of these high-level months are significant. Some of the other

high means when compared with the June low-level provide fairly significant differences. The mean difference between the tallest group and the shortest is just under a third of an inch (.31 inch). These monthly

TABLE 2

Distribution of Heights by Birth-Months

Inches	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sep.	Oct.	Nov.	Dec.
50 Plus								1				
51 "								0				
52 "								0				
53 "								0				
54 "						1		0				
55 "						0		0				
56 "						0		1				1
57 "					1	0	1	0		1		0
58 "			1	2	1	0	2	0		0		0
59 "	1	1	0	0	1	0	2	2	2	4		0
60 "	4	5	4	4	4	4	6	2	4	5	2	1
61 "	11	4	5	6	8	2	5	5	8	6	5	12
62 "	27	20	21	18	20	35	20	16	23	29	20	17
63 "	38	43	64	46	54	56	45	36	46	53	45	50
64 "	92	68	93	96	111	113	98	98	87	105	83	99
65 "	152	143	182	171	158	156	139	161	161	153	139	178
66 "	218	218	229	259	230	273	251	245	250	236	206	219
67 "	251	263	315	262	281	262	306	248	294	277	254	244
68 "	306	233	266	265	276	307	293	267	297	296	272	289
69 "	250	226	251	251	229	226	257	246	244	236	225	253
70 "	171	190	191	178	184	169	181	151	215	193	172	186
71 "	109	108	117	99	130	100	112	116	118	127	103	110
72 "	59	69	62	60	68	57	64	73	75	81	53	78
73 "	43	24	28	33	25	23	26	32	41	42	47	32
74 "	12	15	12	10	14	12	12	13	15	10	25	13
75 "	4	4	6	3	3	0	8	6	2	5	7	2
76 "	3		1		3	2	3	1	4	2	1	1
77 "	0		1		0	0	1					
78 "	1				1	1						

fluctuations are illustrated in figure 1. The main trends of this analysis correspond to those of studies reported by the author in 1941 with reference to the heights of children.

The corresponding facts for weights according to month of birth are given in tables 3 and 4 and figure 1. The distributions in table 4 work

TABLE 3

Weights by Birth-Months

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sep.	Oct.	Nov.	Dec.
Mean (lbs)	150.41	150.11	149.81	149.78	149.48	149.44	149.91	150.41	150.57	149.67	150.19	150.76
P.E.*m*	.3182	.3275	.3061	.3163	.3011	.3066	.3216	.3178	.2946	.3045	.3174	.3155
number	1758	1583	1822	1739	1799	1799	1825	1724	1849	1860	1638	1781

TABLE 4

Distribution of Weights by Birth-Months

Pounds	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sep.	Oct.	Nov.	Dec.
86 Plus						1		1				1
94 "	4	1		3	1	3	6	0	3	2	2	2
102 "	8	4	11	14	4	5	12	8	9	14	5	8
110 "	40	27	34	31	44	50	26	39	34	45	28	34
118 "	78	77	82	88	94	90	87	82	67	85	87	95
126 "	216	208	254	232	248	227	256	204	223	245	211	220
134 "	304	280	323	283	270	307	283	286	282	278	262	276
142 "	300	271	312	315	331	304	328	312	390	388	289	332
150 "	270	261	266	292	306	310	313	262	297	273	271	291
158 "	201	164	224	194	184	189	201	201	219	206	178	175
166 "	151	123	133	115	144	137	143	140	140	122	125	129
174 "	81	59	61	59	66	67	56	75	67	76	77	98
182 "	41	55	51	44	50	52	49	52	52	60	52	56
190 "	27	23	29	31	24	26	27	28	24	28	19	27
198 "	12	10	22	18	22	11	12	12	17	14	17	19
206 "	10	7	14	10	4	10	12	11	12	15	6	7
214 "	6	6	3	2	4	5	7	7	6	6	6	5
222 "	6	4	0	2	1	3	1	1	5	1	1	3
230 "	1	2	1	4	1	1	3	2	2	1	0	2
238 "	1	1	0	2	0	1	2	0		1	2	0
246 "	0		1		0		0	1				0
254 "	1		0		1		0					0
262 "			1				0					0
270 "							0					0
278 "							0					1
—							0					
—							0					
—							0					
—							0					
406 "							1					

with 8-pound intervals instead of the original 4-pound intervals. The lightest men are born in June and the heaviest those born in December with a difference of 1.32 pounds. The June measure is part of an extended low-level period and the December one similarly of a high-level period.

The differences between the lowest weight and the others are not highly significant. This is due partly at least to the skewing in the weight-distributions well indicated by table 4, a fact that is related to the tendency of the older men toward excessive weight. This feature in adult weight was noted by Kendall ('43).

SUMMARY

1. An analysis was made of the height and weight measurements of over 21,000 men of the New Zealand army draft for World War II.

2. The tallest men were those born in February (summer) and the shortest those born in June (winter). There was an average difference of about one-third of an inch between them.

3. The heaviest men were those of December birth and the lightest those born in June. There was an average difference of about one and one-third pounds between these weight-levels.

4. The main differences in height were significant and those in the weights had a lower level of significance.

LITERATURE CITED

Chenoweth, Laurence B. and Richard G. Channing 1941 Relation of season of birth to certain attributes of students, Human Biology, *13*: 533-540.

Fitt, A. B. 1941 Seasonal Influence on Growth, Function and Inheritance, New Zealand Council for Educational Research and Oxford University Press.

Kendall, M. G. 1943 Advanced Theory of Statistics, Vol. I. Chas. Griffin, London.

Mills, C. A. 1941 Mental and physical development as influenced by season of conception, Human Biology, *13*: 378-389.

BOOK REVIEWS

Standard Values in Nutrition and Metabolism. Edited by ERRETT C. ALBRITTON. Pages xiii + 380, W. B. Saunders Co., Philadelphia, 1954. $6.50.

In view of the rapid, almost terrifying rate of accumulation of scientific data, largely quantitative in nature, a periodical critical sifting and compilation of reported results is an important supplement of research productive of new facts. Consequently, the idea of a comprehensive handbook of biological data has a real appeal.

This ambitious project, receiving substantial federal financial support, was entrusted to a special committee, as an organ of the American Institute of Biological Sciences, affiliated with the National Research Council. In addition to the volume on blood (published in 1952) and the present fascicle, it is planned to cover in separate volumes the fields of growth and reproduction, animal and plant physiology, biochemical composition, and toxicology. There is to be, also, an abridged handbook, covering all the areas.

The current volume consists of 233 pages of tables, 16 pages of diagrams, some 100 pages of bibliography indicating the source of the tabular entries, and an index. The readers of this journal are likely to be particularly interested in data on man's daily nutrient allowances (bringing out the unresolved differences between the American, Canadian and British figures); functions of chemical elements and vitamins, with signs of deficiency and excess; excretion of the end products of metabolism; and basal metabolism, including the presently proposed standards and a comparison with other norms. The standards commonly used heretofore in America were based on measurements made on subjects unaccustomed to the procedure and were higher, while the British standards were based on the lowest of repeated measurements on trained persons under rigorously basal conditions. As example, the earlier American (Boothby, Berkson, and Dunn, 1936), current American (Boothby, 1952), and the British (Robertson and Reid, 1952) standards for 25-year-old men are 40.3, 38.4 and 37.1 Cal/sqm/hr, respectively.

It is hoped that anthropometric data, important for the characterization of man's nutriture, will be included in one of the subsequent volumes.

JOSEF BROŽEK

Coronary Heart Disease in Young Adults. By MENARD M. GERTLER, PAUL D. WHITE, AND OTHERS. Pages xviii + 218, Harvard University Press, Cambridge, Massachusetts, 1954. $5.00.

Sanitation and the antibiotics are so successfully solving the health problems caused by other organisms that soon only the faults of the human organism itself may remain to plague us. That these, however, will be more than enough to engage all our research energy is obvious from a cursory glance at our vital statistics. In spite of great advances in diagnosis, surgery and medical care, it is obvious that there is as yet little progress towards the control and prevention of our self-generating ills that now dominate the health picture in this country.

From the standpoint of mortality, the outstanding problem is now coronary heart disease, which bids fair to account for half the deaths in the near future. The appalling frequency of coronary heart disease, however, is perhaps less significant than the fact that it attacks and kills some of us at an early age while some of us seem to be immune. Of course, we want to know why, but the first step towards answering this must be to discover who. In other words, how do the patients with coronary heart disease differ from their more fortunate fellows?

Fifteen years ago Dr. Paul D. White began to attack this question by comparing patients with clinically healthy controls, and the present volume is a report on a continuation of this inquiry with a hundred young (under 40) coronary patients and a suitable control series.

In many characteristics, including relative body weight, the coronary patients were not peculiar, but several significant differences were found. The coronary patients tended to be "stocky" or "mesomorphic," to have relatively high concentrations of cholesterol and of uric acid in the blood, and to produce saliva of rather special characteristics. These peculiarities are quite non-specific, but they are statistically significant.

Much has been made in the past of the fact that men are more prone than women to early coronary disease. In the present study, Gertler, White and their colleagues used various methods to explore the question of a masculinity factor, but the results are inconclusive. They found no difference between the patients and the controls in the excretion of male sex hormones. In physique, the patients tended to be more masculine than the controls but psychologically they were judged to be more feminine.

A dietary inquiry was also inconclusive except that it was clear that

the coronary patients did not consume large amounts of cholesterol nor eat a very heavy diet at the time they were studied. However, this was the situation *after* the development of coronary disease and it would be more useful to know their lifelong dietary pattern *before* the development of the disease and all the resulting changes in the mode of life.

The work by Gertler, White and their colleagues was frankly exploratory. This clear presentation of their findings will help point the way to profitable new studies. By the same token their results tend to discount the probable value of further studies with some of the methods they tried. In sum, this book gives emphasis to the great need for larger and more systematic efforts to discover the characteristics of body and of mode of life associated with the development of atherosclerosis and coronary heart disease. It is probable that there is no single cause but only influences to be found and that all of us are susceptible in some degree. To discover the effective interplay of nature versus nurture will require attack from many approaches and with many methods. Future multidisciplinary research will build on these studies.

ANCEL KEYS

1951 Impairment Study. By the SOCIETY OF ACTUARIES. Pages 300, Society of Actuaries, New York, 1954. $7.50.

For some fifty years, the Society of Actuaries has been publishing in irregular intervals the results of medical impairment studies conducted upon material contributed by American insurance companies. The last major study concerned itself with a twenty-year period ending in 1928. Minor studies were conducted in 1936 and 1938. The present investigation was based upon the experience of 1935-1949 upon the lives of applicants who were medically examined and applied for ordinary insurance policies; 27 life insurance companies representing 70% of ordinary insurance in force in the United States and Canada agreed to take part in this investigation.

The purpose of the study was to make recent mortality experience available for analysis of the effects of specified "impairments" (such as overweight, family history, heart murmurs and other circulatory conditions) on death rates. The practical end was to provide a basis for adjustments in insurance premiums (extra premiums) to be paid by individuals with less favorable life expectancy.

As a control and a reference point, a basic mortality table was constructed reflecting the death rate for the corresponding period among those applicants who were considered unimpaired. Since in previous studies other basic mortality tables have been used, a comparison of the effects of a particular impairment in two or more periods is difficult. For example, if an impairment would show 150% of standard mortality currently and was rated the same in a previous study, it would not necessarily indicate that this impairment has produced comparable mortality in the two periods. It would merely mean that it produced 50% more mortality than the group which, in the corresponding period, was considered standard by the underwriters.

In interpreting the findings for purposes other than those for which they were originally designed, the limitations of the data must be considered. The statistics are derived from a selected sample drawn from among white male adults in the middle income group. It is not necessarily representative of the population as a whole. Mortality ratios were computed on the basis of the number of policies and not on the number of lives. That means that if a person held several policies, he appears statistically like several exposures, and if such person died, several deaths were recorded. While the control group was selected on the basis of a medical examination, the standards were probably not the same as a researcher would wish to apply in selecting a group of clinically healthy persons. The study did not concern itself with the uncovering of more basic associations but mainly with impairment classifications and indicators, some of which do not represent any specific clinical diagnosis but only reports of physical findings or laboratory analysis.

All of this does not mean that the study is not highly significant for the purposes for which it was conducted but that in its present form it is not particularly useful for the practicing physician or the medical investigator. This is not the fault of the authors, who have fulfilled their assignment rather well. It is the result of an unfortunate lack of agreement upon standards between the clinician, the researcher, and the interpreter of insurance medical findings, and of the practical limitations of medical examinations conducted in connection with the issuance of life insurance policies. In some instances, only very recent research work, conducted outside the framework of actuarial mortality investigations, has pointed to significant indicators of "impairment," in the sense of predictors of unfavorable health or longevity (*cf.*, smoking or, apparently, high fat intake and elevated level of blood cholesterol).

In the impairment studies of the Society of Actuaries, there is much material remaining dormant which is important to human biology. Additional criteria of impairment may usefully be added. The insurance industry has a unique opportunity to aid research because it is able to observe large groups of initially clinically healthy persons over long periods of time and evaluate the role of individual differences in "constitution" (with its anthropometric, biochemical, physiological, and psychological aspects) and in the mode of life.

E. H. KLEPETAR

America's Resources of Specialized Talent. The Report of the Commission on Human Resources and Advanced Training, Prepared by DAEL WOLFLE, Pages xviii + 332, Harper & Bros., New York, 1954, $4.00.

This book had its beginnings in the concern of the National Research Council, the American Council of Learned Societies, the American Council on Education, and the Social Science Research Council with the relation between the need for and the availability of highly trained specialists in many areas affecting national welfare and security. The report is the outgrowth of studies that began in 1949 with the appointment of the Commission, with Dael Wolfle as director of research. The work is premised on the assumption that this country's place in world affairs for years to come will depend upon effective utilization of its intellectual resources, and this leads directly into a series of fact-packed chapters that contribute answers, in some measure, to such questions as: What are the prospects for effective utilization of specialized talent? What are the future needs for it? What is the potential supply? How adequate are we utilizing these potentialities?

The report begins with a detailed classification of the fields of specialization. In general the Commission confines itself to those fields for which four or more years of college training are required. This leads into a comprehensive analysis of trends in college graduation by specialized fields, at both the bachelor's and the Ph. D. level. Another basic chapter is devoted to the occupational distribution of college graduates, with particular interest centering in the relationship between the field of college specialty and actual employment. This poses a fundamental question: To what extent is there flexibility in the utilization

of trained manpower? The relationship between degree of specialization and ultimate use of trained ability has significant social implications.

The relation between supply and demand in the specialized fields is carefully analyzed, and as in all of the chapters, the assumptions that underlie the handling of the basic data are carefully stated. By general fields, present size of the professionally employed groups, estimates of current demand, and supply-demand prospects are set forth.

The logical question then follows: How much scientific talent is not being utilized; how much potential ability is being wasted? There is elaborate consideration of who goes to college and the ability-levels of the college population. The variations in intellectual ability among the specialized fields, by educational level, are striking evidence of a selective process that is going on.

Finally, the report discusses the utilization of the educated specialist and considers means for improving utilization of the potential supply.

Certain broad conclusions to which the Commission comes can be suggested here: the trend toward further specialization is unquestionable, and no evidence exists that a saturation point has been or is about to be reached. Because of the close relationship between college graduation and specialization, the future number of such graduates (now about 270,000; in 1970 to be about 590,000) is the key to future supply. Today, known shortages of specialists exist, and some degree of shortage can be expected to exist for many years, at least in some fields. While college graduates tend to remain in their field of original training, there is a large amount of shifting, and this flexibility is "a major national asset," since it contributes to more effective utilization of skilled manpower.

It has further to be concluded, the report demonstrates, that much potentially high talent is being wasted. Colleges could be double in size in this country with no lowering in quality. It is not that the college population is too big; rather, that a significant percentage of high ability youth do not go to college. Fields of specialization do select on an ability basis: students in physical sciences, business and home economics. These differences, overlapping as they are, are less marked than the differences between college and non-college groups.

The Commission believes there is waste in utilizing available trained ability and that there are steps that could be taken to improve this situation. There is need, too, for more attention to identifying, encouraging and even subsidizing the potentially able students so that

they will receive the specialized training that existing and future needs call for.

No review as brief as this can do more than suggest the wealth of material assembled in this volume, and the ingenuity and care with which it has been drawn together and analyzed. Certainly the book does not answer all the questions that it raises, but it provides the best available starting point for the further investigations that should follow. It is a stimulating volume, not only because of the information it provides about existing human resources, field by field, but even more because of its broad implications that fundamentally touch national policy, and through policy, future national security.

MALCOLM M. WILLEY

Human Limbs and their Substitutes. Prepared under the sponsorship of the Advisory Committee on Artificial Limbs, National Research Council. Edited by PAUL E. KLOPSTEG and PHILIP D. WILSON, Pages 714, McGraw, Hill, New York, 1954, $12.00.

Three facts stand out as one attempts to assess the worth and significance of this large volume: 1) the range and diversity of the fields from which the contributors have been drawn; 2) the monumental nature of the work upon a subject which has previously had virtually no modern literature; and 3) the testimony it bears to the admirable coherence of a program which has been centralized about a pressing human problem but which requires for its solution the concerted action of diverse specialists.

The variety of contributing fields to the practice of prosthetics (limb fitting) is explained when it is realized that the stump is produced by the medical man, the artificial limb is fabricated by a paramedical specialist, the prosthetist, using mechanical components which are devised and manufactured by engineering and industry, and the amputee is aided in his biomechanical adjustment to the limb by the physical or occupational therapist, and in his emotional readjustment by the psychologist. Accordingly, there are chapters dealing with surgery, such as amputation, amputee pain and cineplasty; with prosthetics, such as prescription and limb fabrication; and with rehabilitation, such as training, therapy and psychological treatment.

Since it is also a report on recent research and development, the book contains much new and unique material in both the biological and technical categories. Exhaustive locomotion and manipulation studies are reported, and the outstanding developments in replacement mechanisms are described. These applications provide many significant biotechnical lessons. The excellent amputee pain studies represent advances in both theory and medical practice. The development of the cosmetic glove itself is a story of achievement in applied high polymer chemistry.

In summary, *Human Limbs* is an essential reference work for all those who work directly with amputees—surgeons, prosthetists, and therapists. But for many others it provides excellent source material on body mechanics, specialized engineering mechanisms and materials, and rehabilitation concepts and methods.

CRAIG L. TAYLOR

A Methodological, Psychiatric and Statistical Study of a Large Swedish Rural Population. By TAGE LARSSON and TORSTEN SJÖGREN. Acta Psychiatrica et Neurologica Scandinavica, Suppl. 89, Pages 250, Ejnar Munksgaard, Copenhagen, 1954, 25 Swedish krona.

This is the first volume of the newest of the Scandinavian population studies. It describes, for a change, not an isolate but a district selected as average and representative of Swedish rural population, thus providing a psychiatric base line for the whole of the country.

The ambition of the study, as seen in the first volume just published, has been the fullest feasibile ascertainment of the incidence of mental illness. The completeness of ascertainment is of utmost importance because the first step in genetic investigation of mental disorders is a comparison of familial concentration with incidence in the general population. This task has been completed satisfactorily; the study has furnished another set of reliable basic data, and this is its worth.

The authors have investigated the incidence of psychiatric disorders and of low grade mental deficiency in an area consisting of 13 parishes comprising two large and a number of small islands off the west coast of Sweden, close to the mainland. The area had a population of about 25,000 and there were some 1300 registered cases of major mental disorder. In order to avoid the well known but rarely heeded deficiencies of ascertainment inherent in the cross-sectional census method, the

authors scrutinized their population through a 45-year period of observation. Their technique had been seasoned by previous Scandinavian field investigations. Most, if not all, of those investigations had been planned as a technical improvement on the population sampling methods initiated by the Munich school of psychiatric genetics. Judged from this viewpoint, this monograph represents a technical peak in the line of some four decades of endeavor to master the problems of sampling for the biological study of mental disorder.

The usability of their "period method" is largely limited to Sweden with her enviable system of registration. The selected period of observation, 1900-1944, covers the whole of development both of modern psychiatry and genetics. Certainly, the attitude toward mental illness was not the same at the turn of the century as it is at the present time. Clinical opinion, as well as diagnostic methods, too, underwent radical changes. Thus some uncertainty as to completeness and uniformity of ascertainment would be anticipated. The authors themselves acknowledge the shortcomings of this longitudinal observation by stating that the most representative figures were obtained from the period of 1920-1940. Aside from the uncertainties arising from the progress of time one may anticipate diagnostic imperfections inherent in any study depending for basic information on perusal of hospital records and field investigation. The authors are aware of that. One wonders if the labors and good will invested into such extensive field study could not be utilized in a more permanent manner, *e. g.*, by setting up a permanent clinic where a regular diagnostic job on a university hospital level could be done. Study of more neutral humain traits could be easily incorporated, and in this the human biologist would eagerly share.

The analysis of excess mortality linked to psychoses deserves attention. It is here calculated in terms of the remaining mean expectation of life. It has been again found that a psychosis which is not connected with frank somatic pathology shortens life significantly, although it does not kill directly. A few figures give a hint on biological forces which may make people move from their homes: about one quarter of the psychotics born in the area were not residing there at the time of the onset of their illness, which appears to be well in excess of the emigration rate. It can be inferred that certain aspects of the pre-morbid personality made easier the decision to migrate. This would tie in with Ödegaard's finding of a higher incidence of schizophrenia in Norwegian-born persons in Minnesota than in the population of Norway.

There is a great deal of thinking, experience and honest labor behind the many tables and the lengthy explanations of the procedure in the first volume of the monograph. The thoroughness of ascertainment of this study is in keeping with the tradition of Scandinavian population studies. The second volume probably will bring more in the way of general conclusions, and will be eagerly awaited.

KAREL PLANANSKY

BOOKS RECEIVED

HUMAN BODY SIZE AND CAPABILITIES IN THE DESIGN AND OPERATION OF VEHICULAR EQUIPMENT. By Ross A. McFarland, Albert Damon, Howard W. Stoudt, Alfred L. Moseley, Jack W. Dunlap, and William A. Hall. xiv + 239 pp. Harvard School of Public Health, Boston, 1953.

INDIAN HEALTH IN ARIZONA: SECOND ANNUAL REPORT OF THE BUREAU OF ETHNIC RESEARCH. By Bertram S. Kraus. ix + 164 pp. University of Arizona, Tucson. 1954. $1.50.

ATLAS OF MEN: A GUIDE AND HANDBOOK ON SOMATOTYPING. By William H. Sheldon. xvi + 357 pp. Harper & Brothers, New York, 1954. $10.00.

STATISTICS IN RESEARCH. By Bernard Ostle. xiv + 487 pp. Iowa State College Press, Ames, Iowa. 1954. $6.95.

THE PSYCHOLOGICAL VARIABLES IN HUMAN CANCER: A SYMPOSIUM. Edited by Joseph A. Gengerelli and Frank J. Kirkner. vi + 135 pp. University of California Press, Berkeley. 1954. $3.00.

HUMAN ENGINEERING GUIDE FOR EQUIPMENT DESIGNERS. By Wesley E. Woodson. 268 pp. University of California Press, Berkeley, 1954. $3.50.

EXPERIMENTAL PSYCHOLOGY: REVISED EDITION. By Robert S. Woodworth and Harold Schlosberg. xi + 948 pp. Henry Holt and Company, New York, 1954. $8.95.

GENETICA MEDICA: PRIMUM SYMPOSIUM INTERNATIONALE. Edited by Luigi Gedda. v + 467 pp. Istituto Gregorio Mendel, Rome. 1954.

ANALYSIS OF DEVELOPMENT. Edited by Benjamin H. Willier, Paul A. Weiss, and Viktor Hamburger. xii + 735 pp. W. B. Saunders Company, Philadelphia, 1955. $15.00.